ISER Working Paper Series

Infant Health, Cognitive Performance and Earnings: Evidence from Inception of the Welfare State in Sweden

Sonia Bhalotra
University of Essex

Martin Karlsson
Lund University
CINCH, University of Duisburg-Essen

Therese Nilsson
Lund University
Research Institute of Industrial Economics (IFN)

Nina Schwarz
University of Duisburg-Essen

No. 2019-05
May 2019

INSTITUTE FOR SOCIAL & ECONOMIC RESEARCH

www.iser.essex.ac.uk

Non-Technical Summary

We study a historical, pioneering intervention that was implemented in the early 1930s in response to a cessation in the decline of infant mortality in Sweden, at a time when its incidence was similar to that in many of today's poor countries. It constituted a significant step in the development of the modern welfare state in Sweden. Trained health workers provided information, support, and monitoring of newborn health through home visits and local clinics, with a particular emphasis on nutrition and sanitation. The principles were similar to those underlying the Nurse Family Partnership programmes in the UK and the US. Similar early childhood and home visiting programmes are increasingly being introduced in developing countries, but there are few systematic evaluations. In earlier work, we showed that the intervention achieved its goal of improving infant health and survival and, moreover, that it lowered the risk of death from chronic disease and raised longevity, likely through reducing inflammation and improving net nutrition.

In this paper we examine dynamic impacts of the improvement in infant health on longer term educational and economic outcomes. As infancy is a period of rapid neurological development, net nutrition (including breastfeeding, clean water, reduced infections) in infancy can influence brain development, creating a biological mechanism for causal effects of infant health on cognition. Individuals carrying an improved cognitive endowment from infancy may make greater investments in education themselves (lower cost of effort), receive reinforcing investments from parents, or compete more effectively for state investments in education. If the intervention-eligible cohorts exhibit higher human capital attainment and, if there is sufficient demand for the acquired skills, we may expect they had higher earnings.

Programme coverage was universal in treated parishes, for children aged 0--12 months at any time in the window for which the programme was available, for durations that varied with their exact date of birth. We identify causal intent-to-treat effects of the programme using information on birth parish and exact birth date, and allowing impacts to vary with duration of exposure.

We purposively digitised individual birth certificate data from historical parish records, including births before, during and after the trial, to obtain a census of about 25,000 births during 1930--1934 in 114 rural parishes and 4 cities that were representative of the country in 1930. We matched birth registers to administrative school records, the 1970 census files and official tax registers. This results in a unique longitudinal data set in which individuals are tracked from birth to death, while being observed at multiple stages of their life course.

Our main findings are as follows. Primary school test scores improved for programme-eligible boys and girls, although with a markedly different distribution of gains. Boys showed improvements across the distribution, while girls showed larger improvements albeit only towards the upper fourth of the distribution. The distributional results potentially illuminate a mechanism driving the different labour market outcomes observed for men and women. In this era, primary school attendance was universal by mandate, but only about a fifth of all children progressed into secondary school. Secondary schooling was therefore an important margin for the sample cohorts. Entry to secondary school was competitive and, and we show that the chances of entry increased sharply at the top of the primary school test score distribution. In line with this, we find that intervention exposure is associated with a significant increase in secondary schooling for girls, and no change for boys.

In line with this, we find that all labour market gains from exposure to the childhood programme are restricted to women. The intervention led to women having higher employment and earnings at age 36-40, and higher pension income at age 71. Income gains for women were concentrated towards the top of the distribution, and increases in employment were almost entirely in the public sector, in occupations that rank high in cognitive skills using a task-based occupational classification.

The gender differences in programme impacts on labour market outcomes are not a trace of gender differences in the initial impacts of the programme, as the programme had statistically indistinguishable impacts on infant mortality among girls and boys. We provide evidence suggesting they arise instead from the gendered differences we noted in cognitive performance and educational attainment, alongside rapid growth in labour market opportunities that favoured women.

At the time our sample cohorts were making decisions about higher education and employment, Sweden was experiencing a rapid expansion of the welfare state. This created a disproportionate increase in labour demand in public sector occupations dominated by women, in particular, childcare, health and education. Woman-friendly administrative jobs supporting this growing enterprise also increased. This may have been reinforced by two other channels. First, public childcare may have facilitated the labour supply of married women, allowing them to work and move up the occupational distribution. Second, continuing improvements in child health and survival will have contributed to releasing women from replacement fertility and caring for sick children.

Our results contribute to several strands of the literature. First, we contribute to a small body of evidence demonstrating that early life health interventions impact cognitive attainment. This is of enormous significance. Differences in cognitive skills between individuals tend to emerge early and widen with age. There is an ongoing global learning crisis affecting the developing world as well as poor families in developed countries, with millions of children failing to attain their cognitive potential. Policy makers are actively seeking solutions, but often focusing on pre-school stimulation and neglecting to recognize the potential importance of pre-school health.

Second, our findings contribute to a scarce literature providing evidence that cognitive performance and higher education contribute to earnings. Recent studies have suggested that impacts of pre-school programmes on cognition fade, although impacts on earnings often persist, leading to an increasing focus on the relevance of non-cognitive skills.

Third, we contribute to growing evidence of gender differences in dynamic responses to early life interventions. Our results suggest that realising the full impact of an intervention that increases early life human capital may depend upon demand conditions.

There is growing emphasis on universal health coverage, especially for maternal and child health, but few systematic evaluations of immediate or long run impacts. A growing literature documents impacts of early life health on earnings in adulthood but there remains limited evidence of the mechanisms involved and, in particular, the importance of (endogenous) improvements in cognitive performance and educational attainment earlier in the lifecourse. In contrast to many previous studies, we investigate distributional effects (for test scores and income), and this contributes to understanding the mechanisms that drive our findings.

Infant Health, Cognitive Performance and Earnings:

Evidence from Inception of the Welfare State in Sweden

Sonia Bhalotra
University of Essex**

Martin Karlsson
Lund University
CINCH, University of Duisburg-Essen*

Therese Nilsson
Lund University†
Research Institute of Industrial Economics (IFN)

Nina Schwarz
University of Duisburg-Essen ‡

Abstract

We identify earnings impacts of exposure to an infant health intervention in Sweden, using individual linked administrative data to trace potential mechanisms. Leveraging quasi-random variation in eligibility, we estimate that exposure was associated with higher test scores in primary school for boys and girls, with a different distribution of gains, only girls being more likely to score in the top quintile. Subsequent gains, in secondary schooling, employment, and earnings, are restricted to girls. We argue that the differential gains for women accrued from both skills and opportunities, expansion of the welfare state having created unprecedented employment opportunities for women.

Keywords: Infant health; early life interventions; cognitive skills; education, earnings, occupational choice, programme evaluation; Sweden

JEL classification: I15; I18; H41

April 26, 2019

*University of Duisburg-Essen, Weststadttürme Berliner Platz 6-8, 45127 Essen, Germany, and Department of Economics, Lund University, Box 7082, SE-220 07 Lund, Sweden e-mail: martin.karlsson@uni-due.de

†Corresponding author; Department of Economics, Lund University, Box 7082, SE-220 07 Lund, Sweden, and Research Institute of Industrial Economics (IFN), Box 55665, SE-102 15 Stockholm, Sweden e-mail: therese.nilsson@nek.lu.se

**Department of Economics, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, United Kingdom, e-mail: srbhal@essex.ac.uk

‡University of Duisburg-Essen, Weststadttürme Berliner Platz 6-8, 45127 Essen, Germany, e-mail: nina.schwarz@uni-due.de

Acknowledgements: The authors would like to acknowledge the generous support of Riksbankens Jubileumsfond (The Swedish Foundation for Humanities and Social Sciences) P12-0480:1 and the Center for Economic Demography, Lund University. We acknowledge partial funding from ESRC Grant ES/L009153/1 awarded to the Research Centre for Micro-Social Change at ISER, University of Essex. Johanna Ringkvist, Josefin Kilman and Ines Hußmann provided excellent research assistance. We are grateful to Andreas Bergh, Peter Sandholt Jensen, Gustav Kjellsson, Alessandro Martinello, Cheti Nicoletti, Anton Nilsson, Martin Nordin, Bernhard Schmidpeter and participants of various conferences and seminars for their feedback on earlier versions of this paper.

1 Introduction

Spurred by cessation of infant mortality decline, the Swedish government trialled a postnatal intervention from 1 October 1931 to 30 June 1933. It had universal coverage and provided information, support and monitoring of newborn health, including encouragement of breastfeeding, sanitation, a healthy diet and early detection of infections through home visiting and clinic attendance. The explicit purpose of the intervention was to bring down infant mortality, important in itself and also as a marker for improvements in infant health for those who survive infancy (Bozzoli et al., 2009). In previous work, we establish that the intervention achieved this goal (Bhalotra et al., 2017).

In this paper we examine dynamic impacts of the improvement in infant health on educational and economic outcomes. Infancy is a period of rapid neurological development that is sensitive to net nutrition (achieved through breastfeeding, healthy diet, clean water, reduced infections and inflammation). Nutrition matters for childhood development, it being estimated that about 85% of calorie intake in infancy is used to build brains. Avoidance of infections also matters, as severe or repeated infections in infancy tend to divert nutrients away from neurological development (Finch and Crimmins, 2004; Eppig et al., 2010). Moreover, the release of inflammatory molecules during infections may directly impact the developing brain by changing the expression of genes involved in the development of neurons and the connections between them (Deverman and Patterson, 2009). Thus there are biological mechanisms for causal effects of infant health on cognition. Beyond that, individuals carrying an improved cognitive endowment from infancy may make greater investments in education themselves (lower cost of effort), receive reinforcing investments from parents (Yi et al., 2015; Almond and Currie, 2011; Almond et al., 2017; Bhalotra and Venkataramani, 2013; Adhvaryu and Nyshadham, 2016), or compete more effectively for state investments in education (which, we will argue, played a role in the context we study). If the intervention-eligible cohorts exhibit higher human capital attainment and, if there is sufficient demand for the acquired skills, we may expect them to have higher earnings.

A contribution of this paper is that it uses linked administrative data for a large and representative sample of individuals from birth, through school, and to retirement (or death). It is still unusual to have school test scores linked backwards to birth and forwards to labour market outcomes. These additional data help illuminate the mechanisms running from infant health to earnings and, as highlighted in a recent survey on the long arm of childhood exposures (Almond et al., 2017), this is where the evidence is particularly scarce.

We digitised individual birth certificate data from historical parish records, including births before, during and after the trial, to obtain a census of about 25,000 births during 1930–1934 in 114 rural parishes and 4 cities. We show that our sample was representative of the country in 1930 using a range of economic and demographic variables from the 1930 census. Individuals in the birth registers are identified by first name, last name, exact birth date and place (parish) of birth. Using these identifiers, we link the birth records to administrative school records (digitised from regional archives by us), the 1970 census files and official tax registers. We matched 66% of the birth sample to school data, 86% of the birth sample to the 1970 census files and 65% of the birth sample (91% of survivors) to the tax register. Attrition is partly on account of endogenous survival rates, so we investigate it. We find it is differential by treatment status in the school sample, but not in the census and tax register samples. We nevertheless investigate robustness of all estimates to adjusting for attrition. Match rates were similar for men and women, and attrition adjustments will be done by gender.

In treated parishes, children aged 0–12 months at any time in the window for which the programme was available were eligible, for durations that varied with their exact date of birth. We attempt to identify causal intent-to-treat effects of the programme using information on birth parish and exact birth date, and allowing impacts to vary with duration of exposure. In Bhalotra et al. (2017), we report a suite of checks on the 'first stage' reduction in infant mortality which increase our confidence that the variation we exploit is quasi-experimental.

Our main findings are as follows. Primary school test scores at age 10 improved for programme-eligible boys and girls, although with a markedly different distribution of gains. Boys showed improvements across the distribution, while girls showed larger improvements albeit only towards the upper fourth of the distribution. Post-primary school gains in the available indicators are only seen for women. The intervention led to women being more likely to attend secondary school, to have higher employment and earnings at age 36-40, and higher pension income at age 71.[1] Income gains for women were concentrated towards the top of the distribution, and increases in employment were almost entirely in the public sector, in occupations that rank high in cognitive skills using a task-based occupational classification (Autor et al., 2003). These results are robust to controls for parish-specific trends.

[1]Since earnings for women age 36–40 are potentially sensitive to career disruptions related to child bearing or rearing, we examine pension income at age 71. For our cohorts pensions mirror the best fifteen years in the labour market and thus represent earnings at advanced stages of the career. Again we find treatment impacts for women, and not men.

The identified gender differences in programme impacts on labour market outcomes are not a trace of gender differences in the initial impacts of the programme, as the programme had statistically indistinguishable impacts on infant mortality among girls and boys, or gender differences in utilisation (Bhalotra et al., 2017). We provide evidence suggesting they arise instead from the gendered differences we observed in intervention effects on cognitive performance and educational attainment, alongside rapid growth in labour market opportunities that favoured women. The differentiation of results by gender lends credence to the notion of a causal chain running from earlier to later life outcomes. The rest of this section elaborates the results and delineates our contributions in relation to similar studies.

A year of intervention exposure is associated with a total increase in earnings of 7.3%. Restricting to women, the increase is 19.5%. Estimates of unconditional quantile treatment effects following Firpo et al. (2009) suggest no gains anywhere in the distribution for men, and that income gains for women are concentrated in the upper part of the distribution. The probability that women belong to the top earnings quintile increased by 8 percentage points.[2] Later, we discuss a crude cost-benefit analysis which shows that the earnings gains suggest a very high internal rate of return to the intervention. We shall argue that the large increase in mean earnings for women is plausible given that (i) the increase includes an extensive margin increase in employment, and (ii) the intervention raised the cognitive performance and educational attainment of women, propelling them into high-skilled high-wage occupations. In our discussion of the results, we will provide a crude estimate of the relative contributions of employment increases to earnings increases.

We now detail the employment and occupation results. The probability of full time employment among women increased by 7.6 percentage points, an increase of 20.5%. In the pre-intervention period, 92.5% of men in contrast to 37% of women were employed full time, and programme exposure produced no significant change in employment among men. Women were 5 percentage points (29.4%) more likely to work as managers and professionals (and half of this increase is in the health sector), and 4.4 percentage points (35.5%) more likely to work in accounting, banking and administration (almost all of this increase is as office workers or administrators). These were high-wage occupations with a high share of workers with secondary schooling.

Leveraging the data linkage to measure skill acquisition in the school-going years, we were

[2] Outcome quintiles are defined throughout using as reference the distribution of ineligible individuals.

able to directly investigate the extent to which the intervention led to improved cognitive skills. We find impacts at the mean that are larger and only significant for boys (who had lower baseline scores), but girls exhibit larger increases in GPA than boys in the upper regions of the distribution. The intervention increased the chances that girls score in the top quintile by 12.4 percentage points in contrast to an imprecisely determined 2.75 percentage points for boys. The distributional results potentially illuminate a mechanism driving the different labour market outcomes observed for men and women. In this era, while primary school attendance was universal by mandate (Fredriksson et al., 1971), only about a fifth of all children progressed into secondary school. Secondary schooling was therefore an important margin for the sample cohorts. Entry to secondary school was competitive, and we show that the chances of entry increased sharply at the top of the primary school test score distribution. In line with this we find that a year of intervention exposure is associated with a significant 3.5 percentage point increase in secondary schooling for girls, and no change for boys.[3]

In an attempt to investigate more clearly the role of intervention-led increases in skills as a mechanism or mediator determining intervention-led improvements in earnings, we investigate the overlap in intervention-related attainments sequenced over the lifecourse. We find that individuals who had a primary school GPA in the top quintile overlapped with individuals who completed secondary schooling, and each of these groups overlapped with individuals who entered high-ranking occupations and those who experienced an increase in the probability of being in the top quintile of earners. With the caveat that it is only descriptive, a decomposition-style exercise following Gelbach (2016) suggests that secondary schooling was a critical lever, linking test score gains arising from the infant health intervention to earnings increases.

Thus the evidence indicates that one reason that eligibility for the infant health intervention led to women doing better than men is skills. We propose that an additional reason is likely to be that they faced more favourable labour market conditions than men. At the time our sample cohorts were making decisions about higher education and employment, Sweden was experiencing a rapid expansion of the welfare state. This created a disproportionate increase in labour demand in public sector occupations dominated by women, in particular, childcare, health and education. Woman-friendly administrative jobs supporting this growing enterprise

[3]The differential entry of girls was consistent with market returns. For our sample cohorts, Mincerian estimates show that earnings returns to secondary schooling were substantial, and larger for girls.

also increased.[4]

Our results contribute to the literature in the following ways. First, we contribute to evidence demonstrating that early life health interventions have a causal impact on cognitive attainment. Figlio et al. (2014), for instance, state, *"While we have strong evidence from twin comparison studies that poor initial health conveys a disadvantage in adulthood, we have little information about the potential roles for policy interventions in ameliorating this disadvantage during childhood"*.[5]

Second, our findings contribute to a scarce literature providing evidence that cognitive performance and higher education contribute to earnings. Pre-school programmes such as Project STAR and the Perry intervention appear to have raised long term earnings by generating sustained improvements not in cognitive skills but, instead, in health and non-cognitive skills (Chetty et al., 2011; Heckman et al., 2013, 2006; Baker et al., 2018).[6]

Third, we contribute to growing evidence of gender differences in dynamic responses to early life interventions.[7] We argue that differences in both skill accumulation and opportunities mattered. Our finding of greater skill accumulation among girls is consistent with the theoretical framework of Pitt et al. (2012), premised on men having a comparative advantage in brawn-intensive tasks and women in tasks that are relatively intensive in cognitive function. They also observe that gendered responses will depend on gender-specific labour market returns to

[4]This may have been reinforced by two supply-side channels. First, public childcare may have facilitated the labour supply of married women, allowing them to work and move up the occupational distribution (Datta Gupta et al., 2006; Bergh, 2009; Stanfors, 2003). We investigate this later using municipality variation in the childcare expansion. Second, continuing improvements in child health and survival may have contributed to liberating women into the labour market, away from replacement fertility and caring for sick children (Bhalotra et al., 2018).

[5]Figlio et al. (2014) and Black et al. (2007) use sibling or twin estimators to identify impacts of birth weight on cognitive performance in Norway and Florida respectively. Like Figlio et al. (2014), we are able to assess impacts of infant health on cognitive scores at different ages and by the socio-economic characteristics of parents. However, while they analyse impacts of birth weight differences, we use population-level exposure to an intervention that improved infant health. Amongst other studies, Chay et al. (2009) study black-white convergence in test scores as a function of hospital de-segregation in America, Bharadwaj et al. (2013) show impacts of neonatal care facilities on school test scores in Chile and Norway, Bhalotra and Venkataramani (2013) demonstrate impacts of infant exposure to a clean water programme in Mexico on cognitive attainment in middle and late adolescence, and Almond et al. (2009) show that in utero exposure to (accidental) radiation from the Chernobyl disaster influenced cognition. Also see Heckman et al. (2014). Our data contain records of sickness-related absence from school, allowing us to say something, albeit crude, about the relevance of contemporaneous health vs early life health impacts in producing test score gains.

[6]A vast body of research in economics and biology documents long run benefits of early life health interventions on earnings (Almond and Currie, 2011; Heckman et al., 2014; Falk and Kosse, 2016; Bütikofer et al., 2019; Bhalotra and Venkataramani, 2013). While it is implicit that the intervening mechanism is human capital accumulation, there is fairly limited evidence of the importance of cognitive skills in this process. A reason for this is that few previous studies have been able to link data on test scores and educational choices to earnings in adulthood. The seminal study in this domain is Black et al. (2007) who use twin-comparisons to show that IQ and earnings in adulthood are both increasing in birth weight. They do not have any intermediate outcome data.

[7]Studies that find larger educational gains for girls from early life health interventions include Baird et al. (2016); Bhalotra and Venkataramani (2013); Garcia et al. (2018); Molina (2016); Bobonis et al. (2006); Maluccio et al. (2009); Maccini and Yang (2009); Field et al. (2009)

human capital. We demonstrate that the intervention raised the employment of women in skill-intensive public sector occupations, which experienced rapid growth as the welfare state expanded.[8] In a recent review of the literature, Almond et al. (2017) argue that the effects of the early life environment on long run outcomes are often heterogeneous, *"reflecting differences in child endowments, budget constraints, and production technologies"*. Here, we additionally highlight the potential role of opportunities.

Our results are of relevance today. Sweden in the 1930s had an infectious disease environment similar to that in many developing countries today. In principle, modern medicine has progressed but, in practice, the WHO estimates that only a fifth of children who need antibiotics receive them. In any case, the importance of prevention and the role of information concerning diet and hygiene are likely to be similar. There is growing emphasis on universal health coverage, especially for maternal and child health (Gorna et al., 2015), but few systematic evaluations of immediate or long run impacts (Engle et al., 2007). Early childhood programmes similar to the Swedish trial are being introduced in developing countries, for example, the Chilean Crece Contigo Programme (Clarke et al., 2018) and the Indian Integrated Child Development Programme (Dhamija and Gitanjali, 2019), and being refurbished in richer countries, for example, the Nurse Family Partnership in the UK (Cattan et al., 2019).

Our finding that an infant health intervention led to improved cognitive skills is relevant to the stylised fact that differences in cognitive skills between individuals emerge early and widen with age (Flavio and Heckman, 2007; Doyle et al., 2009; World Bank, 2015; Attanasio, 2015). There is an ongoing global learning crisis, with millions of children failing to attain their cognitive potential (UNESCO, 2014). Policy makers are actively seeking solutions, but their focus has been on pre-school stimulation, and there has been more limited recognition of the importance of pre-school health. However, our findings for men also highlight that the average earnings payoff to cognitive skills is uncertain, being dependent upon distributional effects that may determine higher education choices (places in secondary school), and on context-dependent demand conditions (jobs requiring the skills created).[9]

[8] In a similar spirit, Coles and Francesconi (2017) argue that an expansion of job opportunities for women was critical to realisation of the impacts of the pill innovation on women's labour market outcomes in America, and Bhalotra and Venkataramani (2012) show that labour market segregation in the Southern states of America limited realisation of earnings gains from infant exposure to antibiotics for black but not white Americans.

[9] Previous work has discussed changes in the relative demand for female (vs male) labour stemming from recession, war, or technological change (Elsby et al., 2010; Acemoglu et al., 2004; Cortes et al., 2018; Bhalotra et al., 2018), we provide a new perspective, emphasising that expansion of broad-based public services tends to change the relative demand for female labour. This is of potential relevance to understanding prospects for women in developing countries that are currently witnessing large-scale expansion in the provision of schooling, public health services and, potentially, pre-school centres.

The intervention we analyse was a significant pillar in the emergence of the welfare state in Scandinavia. It was positively reviewed by physicians at the time, influencing the roll out of similar nationwide programmes in Denmark and Norway. Our work is related to recent studies examining long-run impacts of these programmes (Hjort et al., 2017; Bütikofer et al., 2019). The Denmark study (Hjort et al., 2017) differs from this paper because it looks only at impacts on adult health. It is instead related to our previous work, which tracked individuals from birth through to death registers identifying age and cause of death and found that exposure to the infant intervention led to reduced adult deaths from infections, cancer and cardiovascular disease (Bhalotra et al., 2017). The Norway study (Bütikofer et al., 2019), emerging in parallel with this study, is more directly related to this paper as it studies years of education and labour market outcomes.

It is important that we broadly reinforce the findings from Norway that an infant health programme led to higher earnings on average. However, there are several differences in context and approach relative to Bütikofer et al. (2019), and the following contributions distinguish this paper. First, we digitised school registers from the 1930s to gain data on individual test scores. It is not uncommon to observe completed years of education, but it is rare to observe test scores and linked them to both an early life intervention, and to future labour market outcomes. Second, we document not just impacts on earnings but also on employment, sector and occupation. This allows us to provide altogether new evidence of the following links in a potential chain of mechanisms: (i) early life health improving school test scores, (ii) test scores as a mechanism linking early life health to later life earnings, (iii) employment and occupation as additional endogenous outcomes that illuminate the pathways running from early life health to higher earnings. Third, we estimate distributional effects for the continuous outcomes, test scores and income, which additionally illuminate mechanisms. Fourth, Bütikofer et al. (2019) (and, similarly, Hjort et al. (2017)) analyse nationwide infant care programmes rolled out over decades while ours was a short pioneering trial announced as concluding within two years, and this limited the possible influence of confounders and unobserved trends. Finally, we identify systematic differences in higher education and labour market performance between men and women, which we trace through the four points of the lifecycle for which we have individual linked data. The story we tell is thus very different.

The rest of the paper is structured as follows: Sections 2 provides background information on the intervention and on the educational system in Sweden in the early 20$^{\text{th}}$ century. Section

3 describes the data and the empirical strategy, while Section 4 presents the results and Section 5 discusses potential mechanisms. Section 6 presents robustness checks.

2 Background

2.1 The Field Trial – Institutional Details

Following declines in maternal and infant mortality at the beginning of the 20$^{\text{th}}$ century, progress stalled during the 1920s, giving rise to an intense public debate in Sweden, and the intervention we analyse emerged as a potential solution. The intervention was described as a trial, implemented prior to a decision on nationwide adoption. It started on 1 October 1931 and ended on 30 June 1933. It was implemented in 7 health districts containing 59 municipalities (2 cities and 57 parishes), namely Lidköping, Hälsingborg, Harad, Råneå, Jokkmokk, Pajala and Mörtfors. They were chosen to be representative of the country in population density and living standards and the selection of districts was not based upon infant or maternal mortality rates, the primary targets of the intervention. The trial was fully funded by the central Government to the tune of SEK 41,400 (USD 139,000 in current prices)(Swedish Government, 1931; SOU, 1935).

To ensure uniform standards of care across the districts, a five-day long educational event for participating staff was organised in Stockholm in July 1931. The trial activities were decentralised to the district level and led by physicians. In each of the seven districts a health centre with regular office hours 2–3 times per week was started. Outreach activities included announcements in local newspapers and churches, and oral announcements by midwives and nurses (Stenhoff, 1934). In total about 2,000 mothers and 2,600 children enrolled.[10] We digitised records maintained by health professionals that indicate programme utilisation. These show that the average infant made 2.8 visits to a health centre and received 3.9 home visits.

The intervention focused on preventive care and included check-ups at surgeries, home visits and information campaigns. Newborn children were weighed and checked, and sick children were referred to doctors. Mothers were encouraged to breastfeed and given written and illustrated details on the nutritional needs of children at different stages of development. Home visits by nurses were designed to provide advice on hygiene, sanitation and cleanliness in the household, and to ensure that families followed guidelines published by the National Board of Health. So as to understand what the control group in our analysis received, it is relevant to note that while

[10]On average 72 per cent of all eligible children in rural areas enrolled in the trial. The share of all eligible children that enrolled in the two urban areas was 52 per cent in Lidköping and 32 per cent in Hälsingborg.

Sweden had a fairly developed primary care system in 1930, there were limited preventive care and support activities targeting infants and expecting mothers.[11]

Eligibility for the infant care programme was determined by birth date. All children less than or equal to 12 months of age at the start of the intervention were eligible and eligibility ceased on their first birthday. Figure 1 shows the duration of eligibility in months for the infant intervention by birth date, the vertical lines representing the beginning and the end of the trial. An antenatal care programme was introduced simultaneously with the postnatal program and all expectant mothers were eligible, irrespective of their stage of pregnancy. The raw correlation of duration of eligibility for the antenatal and the postnatal interventions is 0.32 and, conditional on eligibility for any one intervention, this falls to 0.13. Given the differential exposure of each individual to the antenatal vs the postnatal components of the programme, we can estimate impacts of each conditional on the other. Our estimating equations consistently include a term measuring exposure to the antenatal care programme. However, in Bhalotra et al. (2017), we found no impacts of the antenatal care programme on infant mortality or the later life health of the children and, anticipating the results in this paper, we again find no positive impacts of the antenatal program on the economic outcomes of the births. For this reason, the discussion focuses upon the postnatal (infant care) programme.

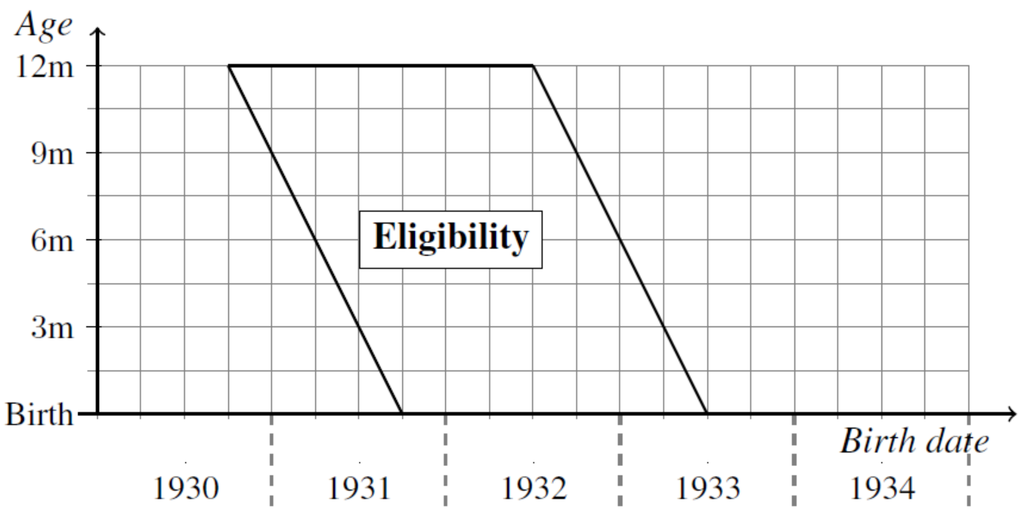


Figure 1. Duration of eligibility by birth date.

[11]The philanthropic childcare institution the *milk drop central* was at the time established in 22 larger cities (among which one city, Hälsingborg, was part of the trial) and engaged in activities to distribute cow milk mixtures to disadvantaged mothers and mothers that could not breastfeed (Wallgren, 1936). The milk drops were generally open twice a week. A Government report of 1929 suggest that they covered around 20 per cent of the infants in the cities where they were established (SOU, 1929). A second type of institution was neonatal care units performing health check-ups and monitoring of mothers and babies. By 1929 there were three active units: in the cities of Stockholm, Gothenburg and Karlskrona (SOU, 1929), cities not included in the trial we analyse.

Annual audit reports in the early 1930s that we perused in libraries indicate programme fidelity, and harmonisation of activities across the treated districts. The trial received positive evaluations from involved physicians in a final report in 1933, attributing improvements in infant health to behavioural change among mothers (Stenhoff, 1934). For example the audit report from the chief physician of the northern districts of Sweden states that there was a notable change in the cleanness and tidiness of childrens' beds and clothing. The first systematic evaluation of the trial is in Bhalotra et al. (2017), where we show that the average duration of programme exposure in infancy led to a 1.56 percentage point decline in the risk of infant death (24% of baseline risk) and a 2.56 percentage point decline in the risk of dying by age 75 (7.0% of baseline risk). We present evidence that intervention-led declines in the risk of dying after the age of 50 were dominated by reductions in mortality from cancer, cardiovascular disease and infections.

2.2 The Swedish School System

In the 1930s, schooling in Sweden started in the year an individual turned seven and was compulsory for six years, and primary education (*Folkskolan*) was universal.[12] Sweden had a tracking system whereby students progressing to secondary schools left *Folkskolan* either after grade 4 or after grade 6. On average barely 20 per cent of children attended secondary school, availability of which increased in the 1940s.[13] Importantly for our purposes, the share of girls and boys attending secondary school was similar in 1930. A reform implemented in 1927 granted equal access for girls to all state-led grammar schools, mandating that girls study the same curriculum as boys. Before 1927 girls could take on secondary education, but only in private schools, the higher costs of which led to lower girl enrollment. By the time our sample cohorts were secondary school age, the situation was transformed, reflecting rapid increases in girls' attendance, particularly in state schools.[14]

Teachers kept records of test scores and attendance in catalogues, which we digitised. The government established several marking principles (see Appendix A). For example guidelines

[12]Parents were legally obliged to send their children to school. Paragraph 51 of the royal decree of the Folkskola stated that parents that did not send their children to school could lose custody.

[13]In the first decades of the twentieth century education beyond the primary level was mainly for children from higher social classes. A series of reforms in the 1925–1945 period, driven by demand and a political will to reduce educational inequalities between urban and rural areas, increased access and the geographical spread of secondary schools. For a discussion on the reforms and the expansions of secondary education, see Lindgren et al. (2014); Kyle and Herrström (1972); Stanfors (2003).

[14]In the late 1920s, among birth cohorts 1915-20, just more than half of all children taking secondary education were in private schools (about 24,000 of 44,000). However, more than 90% of children in state schools were male and more than 90% of children in private schools were female. By about 1940 there was dramatic growth in state provided education, with only 10% of children attending private schools (about 5,000 of 50,000 pupils). Although private schools continued to mostly be populated by girls, there was a gender balance in state schools.

dictated that teachers reward the quality of knowledge and not the quantity, and take notes throughout the year to ensure that grading reflected performance through the year and not at one point in time. The marks we analyse should thus be purged of day-of-test idiosyncrasies. Teachers were instructed to allow for mark inflation as pupils progressed to higher grades, and to make no adjustment for school form. Thus, the marks reflect an absolute standard and not the relative position of a pupil in their class. The test score data are thus fairly reliable. For our sample cohorts, schooling was fairly comparable across districts[15] and the curriculum did not change between 1919 and 1950. The data contain information on school form, a measure of school quality, and we control for this.[16]

3 Data and Empirical Strategy

3.1 Administrative Data Linkage

The dataset is unique in linking individual-level data across the life course using birth registers, school registers, the 1970 census and official tax registers. The birth and school registers were digitised by the authors.

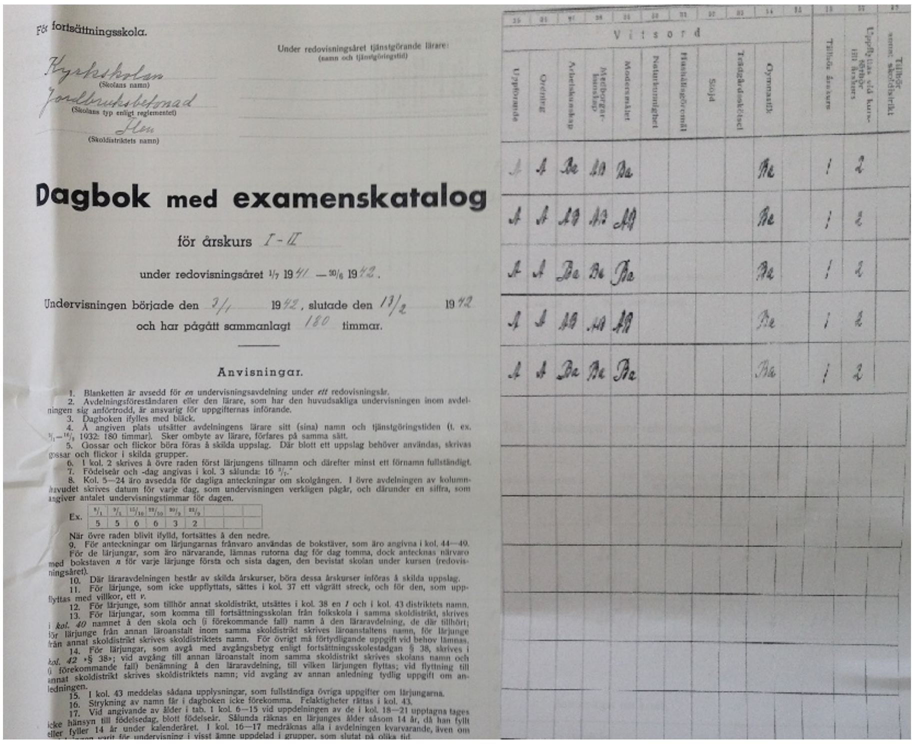
Birth Registers. A census of 24,374 live births in 1930–1934 was digitised from church records, to include births before, during and after the trial of 1931-1933. Sweden is one of the few countries with high-quality vital statistics at the parish level from the 18$^{\text{th}}$ century onwards (Pettersson-Lidbom, 2015). The birth data contain sex, marital status of the mother, age of the mother and parental occupational status, which we translated into occupational classes based on the HISCO classification (Leeuwen et al., 2002) to control for socio-economic status. We merged these birth register data with data from several other sources using linking procedures that were carefully executed and validated; see Bhalotra et al. (2017) for details.

Administrative School Records. We accessed standardised exam catalogues containing pupil-level information from historical archives; see Figure 2. These contain yearly information on school performance and sickness absence in primary school. We observe the birth cohorts of 1930–1934 in grades 1 and 4 of primary school, in school years 1937–1947. Grades 1 and 4 are pivotal as grade 1 represents the first occasion at which school performance can be observed

[15]In 1919 a central education plan, the *utbildningsplanen*, was introduced to overcome differences in the content and format of primary school education across Sweden's 2400 school districts. Guidelines published by the Department of Ecclesiastical Affairs included time-tables, syllabi for compulsory schooling, and a statement of the possible forms a school could have.

[16]Appendix Table J1 provides an overview on the proportion of the school forms in the school year 1940/1941 in comparison to the proportions in our sample.

and grade 4 represents the last as some pupils leave the basic track and proceed to secondary schooling afterwards.



För fortsättningsskola.

Kyrkskolan
(Skolans namn)
Jordbruksbetonad
(Skolans typ enligt reglementet)
Flen
(Skoldistriktets namn)

Under redovisningsåret tjänstgörande lärare:
(namn och tjänstgöringstid)

Dagbok med examenskatalog

för årskurs I - II

under redovisningsåret 1/7 1941 — 30/6 1942.

Undervisningen började den 3/1 1942, slutade den 13/2 1942
och har pågått sammanlagt 180 timmar.

Anvisningar.

1. Blanketten är avsedd för *en* undervisningsavdelning under *ett* redovisningsår.
2. Avdelningsföreståndaren eller den lärare, som har den huvudsakliga undervisningen inom avdelningen sig anförtrodd, är ansvarig för uppgifternas införande.
3. Dagboken ifylles med bläck.
4. Å angiven plats utsätter avdelningens lärare sitt (sina) namn och tjänstgöringstiden (t. ex. 1/1—15/5 1932: 180 timmar). Sker ombyte av lärare, förfares på samma sätt.
5. Gossar och flickor böra föras å skilda uppslag. Där blott ett uppslag behöver användas, skrivas gossar och flickor i skilda grupper.
6. I kol. 2 skrives å övre raden först lärjungens tillnamn och därefter minst ett förnamn fullständigt.
7. Födelseår och -dag angivas i kol. 3 sålunda: 16 3/7.
8. Kol. 5—24 äro avsedda för dagliga anteckningar om skolgången. I övre avdelningen av kolumnhuvudet skrives datum för varje dag, som undervisningen verkligen pågår, och därunder en siffra, som angiver antalet undervisningstimmar för dagen.

Ex.	8/1	9/1	15/10	22/10	30/9	22/9			
	5	5	6	6	3	2			

När övre raden blivit ifylld, fortsättes å den nedre.
9. För anteckningar om lärjungarnas frånvaro användas de bokstäver, som äro angivna i kol. 44—49. För de lärjungar, som äro närvarande, lämnas rutorna dag för dag tomma, dock antecknas närvaro med bokstaven *n* för varje lärjunge första och sista dagen, den bevistat skolan under kursen (redovisningsåret).
10. Där läraravdelningen består av skilda årskurser, böra dessa årskurser införas å skilda uppslag.
11. För lärjunge, som icke uppflyttats, sättes i kol. 37 ett vågrätt streck, och för den, som uppflyttas med villkor, ett *v*.
12. För lärjunge, som tillhör annat skoldistrikt, utsättes i kol. 38 en *1* och i kol. 43 distriktets namn.
13. För lärjungar, som komma till fortsättningsskolan från folkskola i samma skoldistrikt, skrives i *kol. 40* namnet å den skola och (i förekommande fall) namn å den läraravdelning, de där tillhört; för lärjunge från annan läroanstalt inom samma skoldistrikt skrives läroanstaltens namn, för lärjunge från annat skoldistrikt skrives skoldistriktets namn. För övrigt må förtydligande uppgift vid behov lämnas.
14. För lärjungar, som avgå med avgångsbetyg enligt fortsättningsskolestadgan § 38, skrives i *kol. 42* »§ 38»; vid avgång till annan läroanstalt inom samma skoldistrikt skrives skolans namn och (i förekommande fall) benämning å den läraravdelning, till vilken lärjungen flyttas; vid flyttning till annat skoldistrikt skrives skoldistriktets namn; vid avgång av annan anledning tydlig uppgift om anledningen.
15. I kol. 43 meddelas sådana upplysningar, som fullständiga övriga uppgifter om lärjungarna.
16. Strykning av namn får i dagboken icke förekomma. Felaktigheter rättas i kol. 43.
17. Vid angivande av ålder i tab. 1 kol. 6—15 vid uppdelningen av de i kol. 18—21 upptagna tages icke hänsyn till födelsedag, blott födelseår. Sålunda räknas en lärjunges ålder såsom 14 år, då han fyllt eller fyller 14 år under kalenderåret. I kol. 16—17 medräknas alla i avdelningen kvarvarande, även om ...

Vitsord: Uppförande	Ordning	Arbetskunskap	Medborgarkunskap	Modersmålet	Naturkunnighet	Hushållsgöromål	Slöjd	Trädgårdsskötsel	Gymnastik		Tillhör årskurs	Uppflyttas vid kursens slut till årskurs	Tillhör annat skoldistrikt
A	A	Ba	AB	Ba					Ba		1	2	
A	A	AB	AB	AB					Ba		1	2	
A	A	Ba	Ba	Ba					Ba		1	2	
A	A	AB	AB	AB					Ba		1	2	
A	A	Ba	Ba	Ba					Ba		1	2	

Figure 2. Exam catalogue in *Folkskola.*

The data contain academic performance in math, writing and reading and speaking, and religion.[17] Other variables include sickness absence in days, total absence in days, the length of the school year, school type, the name of the teacher and the name of the school.[18] Individuals in the birth records were matched to school records using an algorithm based on birth parish, date of birth, forename and surname. Out of 22,500 individuals still alive at age 7, roughly 16,000 were matched to the school records. It is only for about half of our sample that we have information on both grade 1 and 4, for the rest a child is either observed in grade 1 or in grade 4. Due to the possibility of grade retention there are a few cases where we observe pupils more than once per grade but, at 1.6%, grade retention was rare (Hjalmarsson et al., 2015)[19].

[17] Religion covered Christianity. There were no detailed curricula in the education plan for this subject, only general information about the course content, and no recommended learning material in addition to the bible (SOU, 1946:11). The focus seems to have been bible reading, and learning of psalms and bible texts.

[18] Sickness absence accounts for about 80% of total absence. Other reasons for absence could be inappropriate clothing or weather conditions preventing children from going to school. For details see (Cattan et al., 2017).

[19] The reason for most of the missing information is that the archives of certain schools were accidentally destroyed. Other reasons are death before reaching school age; discrepancies in name spelling; and migration between birth and school age. We significantly reduce the matching problem related to migration by tracking migrants and collecting school records from their destination parishes (parishes not necessarily part of the original sample of treated and matched parishes). Another possible explanation for unmatched individuals could be adoptions since we also match on parental surname, but only about 1% of children in our cohorts got adopted (Bernhardtz, L. and Klintfelt, A., 2007).

Labour Market Outcomes. We merged individuals in the birth records to data from the 1970 population and housing census which covers the entire population of Sweden on 1st November 1970 (Population and Housing Census 1970, 1972a). It contains educational attainment, income, employment status and occupation. Of 24,390 births in 1930-34, we observe 20,922 in 1970. Upon matching birth to death registers, we can see that 3,243 of the 4,142 unmatched individuals died before the 1970 census enumeration.[20,21]

Table 1. Descriptive statistics: Outcome variables.

	Men & Women					Women		Men	
	Count	Mean	SD	Min	Max	Count	Mean	Count	Mean
School Data									
Share Sickness Absence	15,744	0.047	0.054	0	1	7,770	0.049	7,974	0.045
Top GPA	15,789	0.215	0.362	0	1	7,791	0.257	7,998	0.175
GPA	15,789	3.560	0.616	1	6	7,791	3.670	7,998	3.453
Math	15,774	3.518	0.720	1	6	7,780	3.568	7,994	3.469
Reading	15,768	3.618	0.661	1	7	7,779	3.737	7,989	3.503
Writing	14,860	3.519	0.747	1	7	7,341	3.681	7,519	3.360
Census 1970									
Secondary Schooling	20,911	0.182	0.386	0	1	10,298	0.188	10,613	0.176
Working Full time	20,723	0.635	0.481	0	1	10,257	0.336	10,466	0.929
Working Part time	20,723	0.136	0.343	0	1	10,257	0.259	10,466	0.016
Top income	20,921	0.205	0.404	0	1	10,302	0.204	10,619	0.206
log Income	20,921	9.545	1.132	0	13	10,302	8.865	10,619	10.205
Municipal Employment	20,723	0.160	0.367	0	1	10,257	0.232	10,466	0.091
Governmental Employment	20,723	0.085	0.279	0	1	10,257	0.048	10,466	0.121
Tax Registers									
Log Pension Age 71	15,965	11.875	0.448	9	15	8,285	11.704	7,680	12.059

Note: Variable descriptions to this table are available in Appendix B.

Pension Income. We linked the birth records to pension (labour) income available for 2001–2005 from official tax registers. These contain information on 16,194 individuals from the birth records (7,290 individuals having died before age 71 and 1,580 individuals unmatched). An advantage of using pension income is that it is insensitive to career interruptions such as those associated with childbearing, which could influence income observed in 1970 at a prime working age. For the sample cohorts, obtaining a full pension required thirty years of contributions and the level of the pension was based upon the best fifteen years (Sundén, 2006). Table 1 and Appendix Table J2 present descriptive statistics on all explanatory and outcome variables.

Longitudinal Individual Data: Four points in the lifecycle. To summarise, after linking the above datasets, we track outcomes at four different points in the lifecycle. The

[20] We are consequently left with about 900 individuals who cannot be matched. It is possible that they emigrated.

[21] The earnings information in the 1970 census is regarded to be of high quality, but women who were the partners of a small business owner or a farmer could be recorded as working full-time or part-time while having zero taxable earnings. Since this measurement error might bias our results, we impute incomes of these 2,987 women based on their qualifications and hours worked.

potentially treated cohorts are born 1931–1933, and observed in first grade between the school years 1938–1940 when they are 7–9 years old, and in fourth grade between school years 1941–1943 when they are 10–12 years old. We then observe them in 1970 when they are age 37–39, a labour market active age. We match 66% of the birth sample to school data and 86% of the birth sample to the 1970 census files. Appendix Table H9 provides attrition rates by subgroups of eligibility. We observe pension incomes in 2002–2004 when the individuals are 71-73 years old for 91% of survivors (65% of the birth sample). We checked that the match rates were similar for men and women, despite women changing surname at marriage, this being largely because date and parish of birth and first name uniquely identified most people. The match rates for men vs women were as follows. School sample: 65% vs 67%, 1970 Census: 84.5% vs 87%, Pensions: 61.3% vs 71.6%, this difference arising because men were more likely than women to die before pension age.

Matched Controls. Since the intervention took place in seven health districts consisting of 59 municipalities (2 cities and 57 rural parishes), we identified as matched controls, 2 cities and 57 rural parishes (belonging to 38 different health districts) using observable parish characteristics from the 1930 census. The best matches (denoted $\mathcal{J}_M(i)$) were identified using the Mahalanobis distance metric; details are in Appendix C, where we also present further tests and descriptive statistics that validate the matches.

Summary statistics for a range of relevant observables suggest that our analysis sample is representative of Sweden. Table 2 shows 1930 census statistics and the standardised difference (Imbens and Woolridge, 2009) between treated districts and the rest of Sweden. It also shows the standardised difference between treated districts and their matched control. The standardised difference indicate balance across groups and validates the matching procedure. The same holds for other pre-intervention characteristics from annual medical reports reported in the lower panel.

Figure 3 visualises the sample areas at the municipality (parish and city) level. To ensure balance among the matching procedure variables, observations from the control group were weighted based on their population size in 1930 relative to the population size of the treated locations they were matched to. On the one hand this reduces potential bias while on the other hand it will slightly reduce the efficiency of our estimates.

Table 2. Observable characteristics of matched and control districts.

	All (1)	**Treated** (2)	**All Controls** (3)	**Std. Dif.** (2) vs. (3)	**Matched** (5)	**Std. Dif.** (2) vs. (5)
Panel A: Matching Characteristics from the 1930 Census.						
Agriculture	0.340	0.324	0.340	-0.040	0.302	0.054
Manufacturing	0.318	0.340	0.318	0.096	0.345	-0.018
Fertile Married Women	0.121	0.101	0.121	-0.135	0.100	0.060
Income	811	839	810	0.042	847	-0.013
Wealth	2,525	2,703	2,521	0.080	2,655	0.022
Urban	0.334	0.439	0.331	0.158	0.437	0.003
Population	6,271,266	258,418	6,004,052		160,987	
Panel B: Other Pre-Intervention Characteristics.						
Live Birth	0.973	0.974			0.979	-0.024
Wedlock	0.836	0.888			0.884	0.008
Infant Mortality	0.055	0.063			0.064	-0.002
Perinatal Mortality	0.030*	0.017			0.021	-0.017
Infectious Disease	0.005*	0.005			0.006	-0.004
Other Causes	0.020*	0.041			0.038	0.011
Maternal Mortality	348.1	417.275			381.785	0.004
Mother's Age	29.45	29.455			29.610	-0.017
Professional, Technical		0.049			0.038	0.037
Administrative, Managerial		0.025			0.016	0.046
Clerical		0.016			0.025	-0.045
Sales Worker		0.029			0.023	0.031
Service Worker		0.022			0.010	0.071
Agricultural		0.297			0.307	-0.015
Production Worker		0.426			0.460	-0.048
Institutional Delivery	0.242	0.335	0.239	0.151	0.273	0.096
Weeks Compulsory Schooling	226.2	223.8	226.3	-0.244	223.7	0.012
Seven Years Compulsory	0.606	0.838	0.598	0.392	0.666	0.287

Panel A presents local characteristics according to the 1930 census, which were used to match treated parishes to control parishes. *Panel B* presents other local characteristics in the year 1930 which were not available in the 1930 census. Whenever possible, these characteristics are compared with the national averages; however * signifies that national and local statistics not directly comparable. '*Std Dif.*' presents the standardised difference (cf. Imbens and Woolridge, 2009). Column (1) shows summary statistics for the whole of Sweden, column (2) for the treated 57 parishes and 2 cities, column (3) for all non-treated parishes and cities in Sweden and column (5) for the matched controls (57 parishes and 2 cities).

3.2 Empirical Strategy

We estimate impacts of the infant health intervention on academic performance in primary school, secondary school completion, adult employment, occupation and earnings. We use a difference-in-differences (DID) strategy to compare outcomes for exposed cohorts in treated regions to unexposed cohorts and control regions. In contrast to the case in most DID designs, our intervention is switched on and off, as a result of which unexposed cohorts include ineligible individuals born before and after the exposed cohorts.

The estimated equation is

$$y_{ipt} = \alpha + \beta T_t + \gamma_p + \tau T_t D_p + \sigma_t + \lambda X + u_{ipt}$$

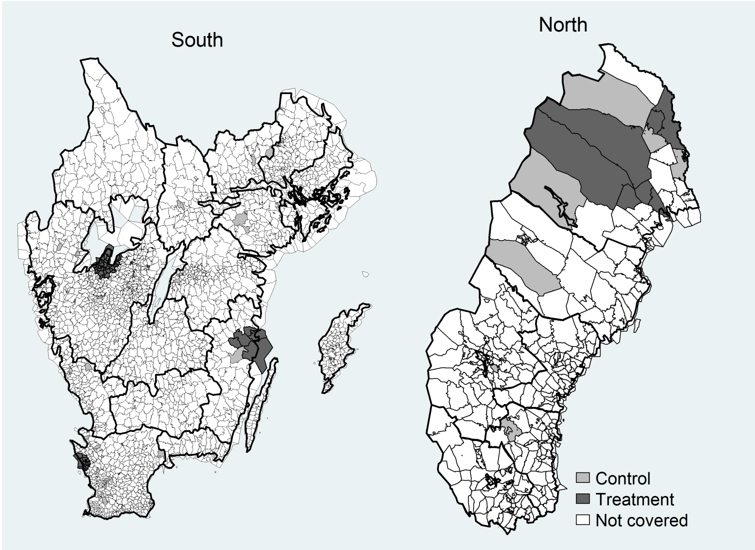


Figure 3. Municipalities containing treated and control districts.

where y_{ipt} is the outcome for child i born in parish p on day t, T_t is the duration of eligibility for the intervention for child i born on day t in years, D_p is a dummy equal to one for treated parishes, γ_p are parish fixed effects, σ_t are $Quarter\ of\ birth \times Year\ of\ birth$ fixed effects and X is a vector of covariates.

Covariates that we condition on include the sex of the child, whether the child was born in a hospital, marital status of the mother, a twin indicator, dummies capturing older (>35 years) and younger (<25) mothers and the occupational status of the household head at the birth of the child. We also control for eligibility for the maternal (prenatal care) intervention since some individuals were eligible for both interventions. The richness of the information in the school records allows us to also control for school fixed effects, length of the school year, and school form (an indicator of school quality). In order to allow for differential trends in outcomes between treatment and control regions, we investigate robustness to including parish specific time-trends, which are more general than treatment-group-specific trends.[22]

The parameter τ measures the intent-to-treat (ITT) effect of the infant intervention for an additional year of eligibility. This is the parameter of interest for policy makers who are unable or unwilling to make the utilisation of services mandatory. Since there were no always-takers (cf. De Chaisemartin, 2012) the ITT is a scaled version of the average treatment effect on the treated (ATT). As in all studies of the long run effects of a positive health intervention, surviving individuals are negatively selected and, as a result, our estimates will be conservative. Indeed,

[22]We also checked that our findings are robust to including health district fixed effects and health district specific trends. Counties contain health districts which consist of parishes, which are in 99% of cases identical to school districts.

we demonstrate this for women.

In Bhalotra et al. (2017) we presented results which increase our confidence that the programme variation across birth cohort and birth parish that generated the initial improvement in infant health is quasi-experimental. We presented evidence that we could reject the concern of differential pre-trends in infant mortality between treated and control parishes. We showed that estimates using within-mother variation in outcomes are similar, suggesting no selection into programme uptake. Using data that we digitised from practitioner records of programme utilisation, we also found no evidence of selective utilisation, by socioeconomic status or gender of the child (see appendix D). We also showed that there were no programme impacts on fertility and that programme impacts on infant and adult health were not sensitive to accounting for trends in hospital birth, parish-specific impacts of the Great Depression, and two school reforms, and that they were robust to randomisation inference.

In Section 6, we present a number of specification checks for the long run outcome estimates in this paper. We use alternative definitions of the treatment indicator. Instead of a continuous indicator which takes duration of eligibility into account, we investigated the effect using binary variables to identify age and duration of exposure. Since marks were given on an ordinal scale, our findings are potentially sensitive to the choice of scale (cf. Bond and Lang, 2013; Lang, 2010; Cunha et al., 2010). To address this we anchor the 7-point grading scale to the logarithm of income in adulthood (as proposed e.g. by Cunha and Heckman, 2008). This is one more advantage of the fact that we are able to link school test scores to earnings at the individual level. We assess the impact of survival selection on the estimates by presenting estimates for sub-samples surviving to different age thresholds.

We use pre-intervention outcome data to formally test for differential pre-trends in the education and labour market outcomes between treatment and control regions. We test if attrition is different by treatment status, and assess sensitivity of our estimates to attrition by estimating Lee bounds (Lee, 2009) for all the main outcomes considered.[23]

We implement a placebo test using a fake intervention ten years after the actual intervention. Finally, we conduct a randomisation inference test for the long-term outcomes, randomly assigning treatment status within each treatment and control parish pair. We then plot the

[23]In estimating the bounds, we exploited the fact that a number of baseline characteristics are available. Thus, the trimming was done conditional on the child's gender, marital status of the mother, and a binary indicator for the household head having a high-SES occupation.

distribution of placebo treatment effects alongside the actual treatment effect.[24]

We investigate the role of school reforms and World War II as confounders. Parliamentary decisions in 1936 and 1937 led to the roll-out of an extension of compulsory school years and of the length of the school year, which all school districts were to have implemented by the late 1940s, see Fischer et al. (2017) and Fischer et al. (2013). The term length extension, which extended the school year by 3–5 weeks (8–13%), would affect students in all school years, and the extension of compulsory schooling from 6 to 7 years affected pupils who did not proceed to secondary schooling. In Bhalotra et al. (2017) we showed that these reforms are largely unrelated to the intervention studied here, but we nevertheless control for both reforms in our analyses. Sweden was neutral during the Second World War and historical sources suggest no educational disruptions for our sample cohorts. In any case, we control for year fixed effects.[25]

In Section 5, we discuss an approach to assessing mediating factors. Having presented estimates for intermediate and final outcomes, we descriptively estimate the extent to which treatment effects that acted on outcomes earlier in the life course contributed to treatment effects on outcomes later in life.

4 Results

We first present results for education and earnings, examining test scores (at age 7 and 10), progression to secondary school (the relevant margin for higher education in the sample period) and earnings (measured when the marginal cohort is 39, and 71). In the next section, we explore the proximate sources of changes in these outcomes by examining sickness absence in school, and employment and occupation in adulthood (age 39).

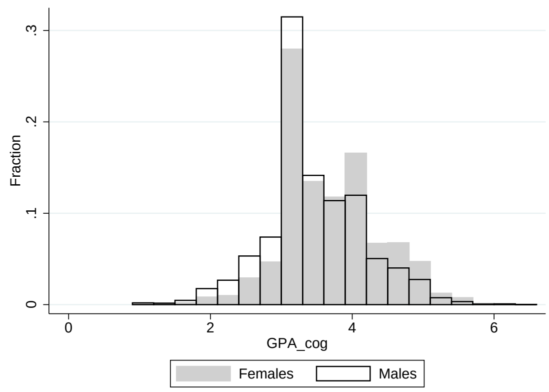
[24] Plotting coefficients for every birth date in an event study style specification is not straightforward because our specification uses duration of exposure during infancy and we expect programme impacts to vary by duration and age of exposure. Also, since the age composition of children varies across the duration of the programme, we cannot read linearity in age from the plot. Figure 4 in (Bhalotra et al., 2017) is an event study graph describing the relationship between infant mortality and programme eligibility by birth quarter. It indicates programme impacts which are roughly proportional to eligibility although the estimates are noisy, in particular towards the end of our sample. The specification checks discussed above serve the purpose more effectively.

[25] In fact the *Folkskola* was one of the main social agents for some 50,000 Finnish children (in ages one to ten) that were evacuated to foster care in Swedish families during World War II. This said, schools were allowed to have shorter breaks in case of limited energy supply, and schools could cancel regular schooling in case of a threat but any lost days had to be replaced by additional days later on, and in case a teacher was called for military service he had to be replaced by a substitute teacher (Fredriksson et al., 1971). We take care of the latter by controlling for school form and we check whether there are any structural breaks in our school data during the war years. We do not find any evidence of disruption in schooling due to the Second World War.

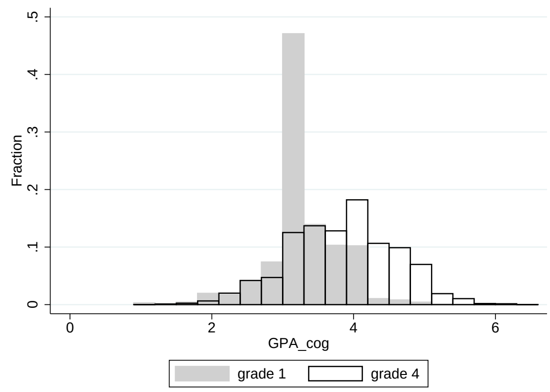
4.1 Outcomes: Human Capital and Earnings

4.1.1 Cognitive Performance- Primary School

As discussed, we digitised school records from paper files, drawing from archives across the country, and matched them to the birth data. We created a measure of cognitive ability by taking the mean of grades in math, reading and speaking and writing to form a grade-point average (GPA), although we shall also report subject-specific estimates. Girls, in general, got better marks than boys, and marks in grade 4 exhibit a higher mean and greater spread than in grade 1. We present results separately for grade 1 and grade 4, and for boys and girls.[26] In order to ease interpretation of the coefficients we transform grades into a z score using the inverse standard normal distribution, and these data are plotted in Figure 4 by gender and grade. [27]



(a) Girls vs. Boys



(b) Grade 1 vs. Grade 4

Figure 4. Distribution of Test Scores.

Table 3 presents our estimates. Exposure to the infant health intervention leads to a statistically significant increase of about 0.08 standard deviations in total GPA in grade 4 (Panel B, row 2). The estimated coefficients are not significantly different by sex, but are larger and only statistically significant among boys, who exhibit a GPA increase of about 0.11 standard deviations. However, when we instead consider the probability of being in the top quintile of GPA, a different pattern emerges (Panel B, row 1). On average, the effect is 7.5 percentage points, and the effect is largely driven by females, for whom it is 12.4 percentage points. These results are robust to controls not only for parish and birth quarter $\times$ year fixed effects but also school fixed effects, the length of the school year and school form, indicators of the socioeconomic status of

[26] The individual correlation between GPA in grade 1 and grade 4 is 0.46, see Figure J1.

[27] As discussed in SOU (1942) the recommendation for teachers was to be restrictive with any high or low marks for children in grade 1 and 2, so a lower variation in the first year of Folkskola as compared to grade 4 is expected. The figure shows the normalized scores.

the parents of the child, and parish specific trends.

Disaggregating GPA by subject, we see that the significant improvements are in 'writing' and 'reading and speaking', which increase by about 0.11 and 0.12 standard deviations on average. These increases are not significantly different by gender, but are larger and only statistically significant for boys. The coefficients for boys are 0.13–0.18 standard deviations, and for girls 0.08–0.11 standard deviations.[28] Panel A of the table shows that there is no discernible impact of the programme in grade 1. As some recent studies have found that cognitive gains stemming from pre-school interventions fade (see e.g. Bitler et al., 2016) and Chetty et al. (2011), while theory predicts that the gains will multiply over time, it is notable that the infant health intervention we consider produced cognitive gains that only become evident at age 10-12. Examining heterogeneity by socioeconomic status we find that children born out of wedlock benefited substantially more than other children but there were no differences in effects by parental socio-economic status (as indicated by their occupation); see Appendix E.

Plots showing unconditional quantile treatment effects, following Firpo et al. (2009), are in Figure 5 for grade 4 GPA by gender. While boys experienced positive treatment effects across most of the distribution, it was only in the upper 30% of the distribution that girls benefited from the intervention. The intervention increased the chances that girls score in the top quintile by 12.4 percentage points in contrast to an imprecisely determined 2.75 percentage points for boys.

To put the average gain in cognitive performance of 0.11 standard deviations in perspective, consider that Bharadwaj et al. (2013) identify effects of 0.15-0.22 s.d. in Chile and Norway using a sample of children at the low birth weight margin. Using twin fixed effects Bharadwaj et al. (2017) estimate that a 10% increase in birth weight in Chile increases outcomes in math and language scores by 0.04-0.06 standard deviations, and examining twin pairs in Florida, Figlio et al. (2014) estimate that, on average, the heavier twin scores about 0.05 s.d. better than the lighter twin. Bhalotra and Venkataramani (2013) find that a 1 s.d. decline in infant exposure to diarrhea following a water chlorination programme in Mexico led to a roughly 0.1 s.d. increase in Raven scores and a 0.07 s.d. increase in math and reading scores. Thus, our estimates are sizable. In fact, they look fairly large even in relation to educational interventions in developing countries, some of which have shown test scores gains between 0.17 s.d. to 0.47

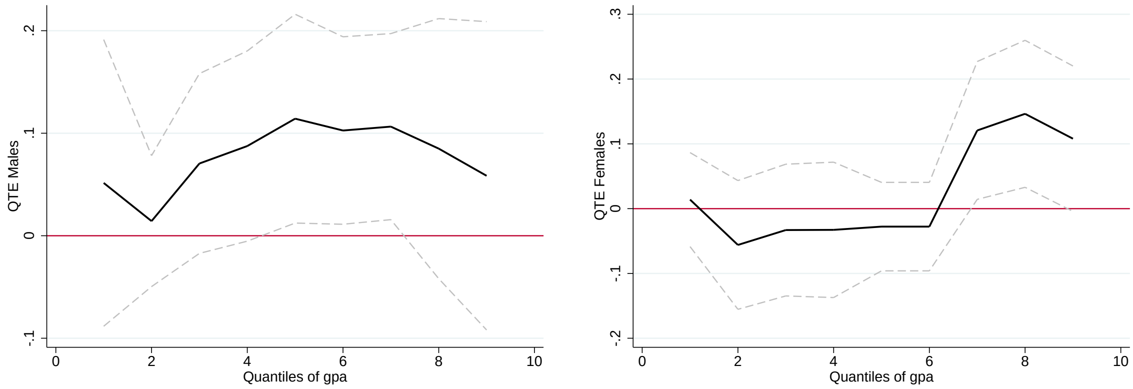
[28] Our finding that effects on literacy dominate effects on math are also seen in some previous work, for example, Sievertsen and Wüst (2017) which estimate effects of same-day post-birth discharge and Aizer et al. (2018), which estimates effects of reduction in blood lead levels in pre-school, but other studies show math scores respond similarly to reading (see e.g. Figlio et al., 2014; Almond et al., 2014; Bhalotra and Venkataramani, 2013).

Table 3. Cognitive Performance – Primary School

	Boys & Girls				Girls				Boys			
	N	Mean	(1)	(2)	N	Mean	(3)	(4)	N	Mean	(5)	(6)
Panel A: Grade 1												
Top GPA	13,207	0.204	-0.0002	-0.0057	6,404	0.223	0.0284	0.0308	6,803	0.185	-0.0193	-0.0302
			(0.035)	(0.038)			(0.050)	(0.057)			(0.033)	(0.033)
GPA	13,207	-0.032	0.0129	-0.0020	6,404	0.027	0.0406	0.0468	6,803	-0.093	-0.0015	-0.0352
			(0.050)	(0.053)			(0.075)	(0.084)			(0.050)	(0.049)
Math	13,161	-0.058	-0.0327	-0.0525	6,382	-0.050	0.0201	0.0014	6,779	-0.066	-0.0895*	-0.1121**
			(0.050)	(0.050)			(0.089)	(0.090)			(0.049)	(0.045)
Reading	13,177	0.001	0.0331	0.0170	6,383	0.082	0.0393	0.0570	6,794	-0.082	0.0534	0.0094
			(0.050)	(0.053)			(0.062)	(0.073)			(0.068)	(0.071)
Writing	9,007	-0.016	0.0937	0.0891	4,399	0.091	0.1469	0.1719	4,608	-0.131	0.0528	0.0225
			(0.093)	(0.093)			(0.126)	(0.129)			(0.074)	(0.075)
Religion	13,060	-0.027	0.0107	-0.0223	6,337	-0.000	0.0031	-0.0131	6,723	-0.054	0.0326	-0.0160
			(0.067)	(0.071)			(0.097)	(0.106)			(0.062)	(0.066)
Panel B: Grade 4												
Top GPA	13,268	0.173	0.0697*	0.0749*	6,561	0.227	0.1000*	0.1243*	6,707	0.116	0.0400	0.0275
			(0.039)	(0.039)			(0.059)	(0.071)			(0.033)	(0.028)
GPA	13,268	-0.047	0.0737**	0.0759**	6,561	0.098	0.0410	0.0617	6,707	-0.200	0.1213**	0.1084
			(0.033)	(0.036)			(0.049)	(0.054)			(0.057)	(0.072)
Math	13,242	-0.027	-0.0220	0.0010	6,554	0.025	-0.0535	-0.0217	6,688	-0.082	0.0193	0.0317
			(0.047)	(0.045)			(0.051)	(0.056)			(0.079)	(0.091)
Reading	13,223	-0.056	0.1179**	0.1105*	6,536	0.120	0.0832	0.0902	6,687	-0.241	0.1823***	0.1649**
			(0.045)	(0.059)			(0.057)	(0.066)			(0.064)	(0.082)
Writing	13,228	-0.057	0.1239**	0.1129**	6,536	0.150	0.0859	0.1068	6,692	-0.275	0.1645**	0.1291*
			(0.056)	(0.054)			(0.081)	(0.094)			(0.064)	(0.072)
Religion	13,238	-0.044	-0.0150	0.0372	6,549	0.088	0.0160	0.0654	6,689	-0.184	-0.0222	0.0247
			(0.049)	(0.035)			(0.052)	(0.066)			(0.097)	(0.096)
Parish FE			✓	✓			✓	✓			✓	✓
QOB×YOB FE			✓	✓			✓	✓			✓	✓
School FE			✓	✓			✓	✓			✓	✓
SES Effects			✓	✓			✓	✓			✓	✓
Length of Schoolyear			✓	✓			✓	✓			✓	✓
Schoolform			✓	✓			✓	✓			✓	✓
Parish Trends				✓				✓				✓

Note: *** p <0,01; ** p <0,05; * p <0,1, Standard errors are clustered at the parish-grade level. Covariates which are included in all specifications are a dummy indicating twin births, a dummy for being female, dummies capturing old (>35 years) and young (<20) mothers, a dummy for married women, a dummy indicating a hospital birth and the treatment effect of the maternal intervention. *Pre-mean* refers to the mean value of the outcome variable before the intervention took place. *QOB× YOB effects* include quarter-of-birth dummies for each of the 20 quarters. *Parish FE* are fixed effects for the parish the individual lived in at the time of the birth. *SES effects* are fixed effects for the professional group of the household head. *Length of schoolyear* are fixed effects controlling for the reforms concerning the length of the school year. *Schoolform* are fixed effects controlling for the school form as described in Section 2.2 and *Parish specific linear trends* allows for parish specific time trends.

s.d. (Duflo and Hanna, 2005; Muralidharan and Sundararaman, 2011; Banerjee et al., 2007), while cash transfer programmes have shown limited impacts, with coefficients ranging between 0.04 to 0.08 s.d. across five studies, and consistently not statistically significant (Baird et al., 2014). Research and policy concerned with improving cognitive attainment has paid increasing attention to the pre-school environment, including parenting styles, caregiver quality and the role of stimulation (Heckman, 2006; Attanasio et al., 2014; World Bank, 2015). Our estimates suggest that pre-school health interventions have the potential to raise cognitive attainment as much as interventions that directly target cognitive capacity.



(a) Boys (b) Girls

Note: Covariates which are included are a dummy indicating twin births, dummies capturing old (>35 years) and young (<20) mothers, a dummy for married women, Parish FE and QOB×YOB FE. 90% Confidence Intervals included.

Figure 5. Quantile regression: GPA in grade 4 by gender.

4.1.2 Secondary Education

The estimates in Table 4 show that an additional year of exposure to the intervention resulted in a 3.5 percentage point (17.6%) increase in the probability that girls completed secondary school, while there was no change among boys. The pre-intervention mean is somewhat larger for girls, but not significantly different from that for boys. The female/male student ratio in secondary education started to increase after a 1927 reform but was close to 1 throughout the 1930's, before increasing further in the 1940–50 period (Schånberg, 1993).

We showed that the intervention led to girls being more likely to score grades in the upper part of the distribution (Figure 5), and baseline performance was already stronger among girls. This is likely to have contributed to the intervention leading to higher secondary schooling increases for girls than for boys– see Figure 6 which shows secondary schooling rates increasing sharply with primary school test scores towards the top of the score distribution.[29]

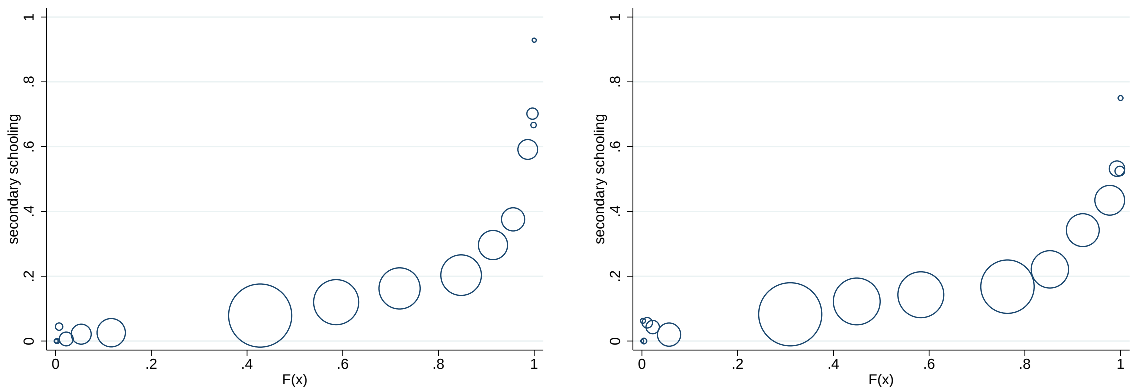
The greater entry of intervention-eligible girls to secondary school is also consistent with higher returns to secondary school among girls. Using the 1970 census, we regressed income in 1970 on test scores in grade 4 and an indicator for completion of secondary schooling (Table 5). We find higher earnings returns to both school grades and secondary schooling among girls than

[29]Students seeking entry to secondary education had to take an entrance test (Wallin and Grimlund, 1933). The test was national, covered certain subjects (Swedish and math, written and oral tests) and only students who passed the test were eligible for secondary schooling. For acceptance, students also needed to pass in other subjects in primary school (Dahr, 1945). Despite an increasing number of secondary schools, there were more applicants than available seats, particularly in urban areas. According to Skolöverstyrelsen (1955), about 11 per cent of all applicants of the cohorts born 1930-1934 were rejected. This may contribute to explaining why the intervention did not raise secondary schooling for boys, even though on average they exhibited higher test scores as a result of the intervention

Table 4. Secondary Schooling

	Boys & Girls ($N = 20,474$)			Girls ($N = 10,105$)			Boys ($N = 10,369$)		
	Mean	(1)	(2)	Mean	(3)	(4)	Mean	(5)	(6)
Primary	0.700	0.0099	0.0180	0.675	-0.0087	-0.0011	0.725	0.0280	0.0367
		(0.021)	(0.020)		(0.032)	(0.026)		(0.023)	(0.025)
Dropout	0.114	-0.0031	-0.0222	0.126	-0.0196	-0.0277	0.101	0.0131	-0.0167
		(0.016)	(0.020)		(0.023)	(0.023)		(0.029)	(0.027)
Secondary	0.185	-0.0062	0.0027	0.198	0.0353**	0.0350**	0.172	-0.0468	-0.0289
		(0.017)	(0.013)		(0.016)	(0.014)		(0.029)	(0.021)
Parish FE		✓	✓		✓	✓		✓	✓
QOB×YOB FE		✓	✓		✓	✓		✓	✓
SES Effects		✓	✓		✓	✓		✓	✓
School Reforms		✓	✓		✓	✓		✓	✓
Parish Trends			✓			✓			✓

Note: *** p <0,01; ** p <0,05; * p <0,1, Standard errors are clustered at the parish level. Covariates which are included in all specifications are a dummy indicating twin births, a dummy for being female, dummies capturing old (>35 years) and young (<20) mothers, a dummy for married women, a dummy indicating a hospital birth and the treatment effect of the maternal intervention. *Mean* refers to the mean value of the outcome variable before the intervention took place. *QOB×YOB effects* include quarter-of-birth dummies for each of the 20 quarters. *Parish FE* are fixed effects for the parish the individual lived in at the time of the birth. *SES effects* are fixed effects for the professional group of the parental household head. *School reforms* refers to the extension of compulsory schooling and length of school year reforms and *Parish trends* allows for parish specific time trends.



(a) Boys (b) Girls

Note: Circle size indicates number of people in each group.

Figure 6. Correlation of marks in primary school and secondary schooling completion.

boys (see also Björklund and Kjellström, 1994).[30] Our findings are in line with the predictions of (Pitt et al., 2012), premised on men having a comparative advantage in brawn-intensive activities, and women in cognition-intensive tasks. Bhalotra and Venkataramani (2013) find broadly similar results in Mexico in the 1990s and Saaritsa and Kaihovaara (2016) in Finland in the early 20$^{\text{th}}$ century, who argue that boys were more likely to drop out of school because of

[30]Previous work suggests that, in the 1930s too, the returns to years of schooling were greater for women than for men (Bång, 2001) and then lifetime returns to education increased for women in particular following a legal reform implemented in 1939 which prohibited firing women on grounds of marriage or pregnancy, similar to the lifting of marriage bars in the United States (Goldin, 1988).

lower net expected returns to schooling, while better educated girls benefited from the expansion of modern services creating attractive working conditions.

Table 5. Returns to education.

	Men & Women (1)	Women (2)	Men (3)
Standardised Grade 4 GPA	0.0841***	0.0905***	0.0798***
	(0.011)	(0.021)	(0.010)
Secondary Schooling	0.4645***	0.5379***	0.3837***
	(0.024)	(0.042)	(0.023)
Female Child	-1.3704***		
	(0.016)		
Constant	10.3144***	8.8962***	10.3662***
	(0.065)	(0.118)	(0.059)
N	12,518	6,221	6,297
R^2	0.385	0.045	0.103

Note: *** p <0,01; ** p <0,05; * p <0,1, Outcome variable is log income 1970. Covariates which are included are a dummy indicating twin births, dummies capturing old (>35 years) and young (<20) mothers, a dummy for married women, and dummies indicating parental SES.

4.1.3 Earnings

We estimate that a year of exposure to the infant health intervention raised earnings by 7.3% on average (Table 6), driven entirely by women experiencing an increase of about 19.5%, in contrast to no gain among men. Unconditional quantile treatment effects for women's earnings show that the increases are concentrated in the upper part of the income distribution (Figure 7), the probability of belonging to the top quintile of earners (the variable *Top Income*) increases by 7 percentage points.

The large earnings increase among women is entirely plausible. First, in the next section, we show that the intervention also raised employment among women, but not men, so there was an extensive margin increase, which we estimate can account for four-fifths of the observed increase.[31] It is useful here to note the vast difference in earnings between women working

[31]Suppose that prior to the intervention, n_2 individuals work full-time, n_1 individuals work part-time and $1 - n_1 - n_2$ individuals do not work. Their log earnings are y_2, y_1 and y_0, respectively. After the intervention, n_2^1 individuals work full-time and n_1^1 individuals work part-time. The extensive margin effect on earnings may then be calculated as

$$\frac{\Delta y}{y_0} = \frac{\left(n_2^1 - n_2\right)\left[\exp\left(y_2\right) - \exp\left(y_0\right)\right] + \left(n_1^1 - n_1\right)\left[\exp\left(y_1\right) - \exp\left(y_0\right)\right]}{n_2 \exp\left(y_2\right) + n_1 \exp\left(y_1\right) + n_0 \exp\left(y_0\right)} \tag{1}$$

In our case, $n_1^1 - n_1 = 0$, $n_2^1 - n_2 = 0.076$, $y_2 = 9.89$, $y_1 = 9.18$, $y_0 = 7.93$. Hence, we get:

$$\frac{\Delta y}{y_0} = \frac{0.076 \cdot 16,953}{8,022} = \frac{1,288}{8,022} = 16\% \tag{2}$$

full-time and women who worked part-time (earnings less than half) and women who did not work (earning one seventh; employment refers to the census week and earnings to the year). Second, we have noted that the intervention led to higher cognitive performance and secondary schooling completion rates. Using a similar approach to that for employment, we estimate that secondary schooling contributed 2.3 percentage points to the increase in earnings, or 1.7% if test scores are held constant.

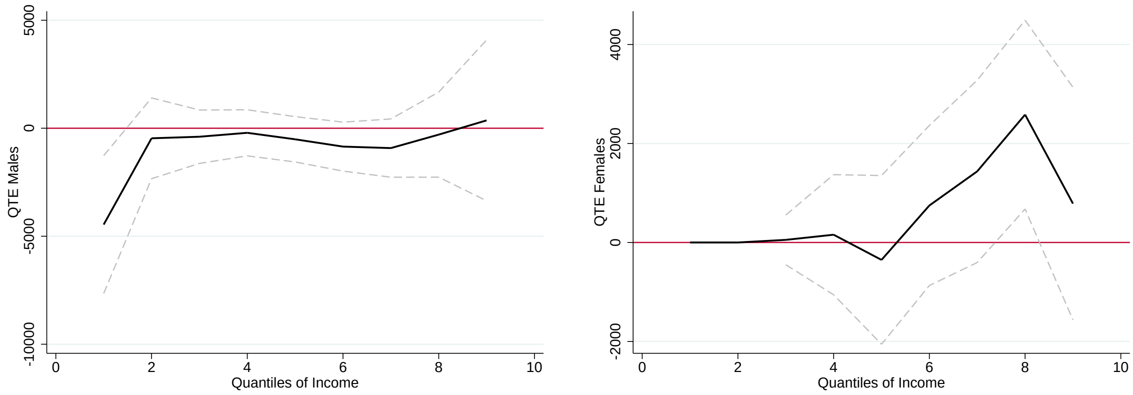
Table 6. Earnings

	Men & Women			Women			Men		
Census 1970	N=20,920			N=10,307			N=10,613		
	Mean	(1)	(2)	Mean	(3)	(4)	Mean	(5)	(6)
Top Income 1970	0.228	0.0099	0.0209	0.244	0.0655***	0.0788***	0.210	-0.0445	-0.0361
		(0.016)	(0.013)		(0.022)	(0.028)		(0.034)	(0.028)
log Income 1970	9.593	0.0295	0.0732**	8.990	0.1204*	0.1947***	10.222	-0.0596	-0.0464
		(0.033)	(0.028)		(0.063)	(0.066)		(0.037)	(0.036)
Pension Age 71	N=15,964			N=8,284			N=7,680		
	Mean	(1)	(2)	Mean	(3)	(4)	Mean	(5)	(6)
log Pension	11.789	-0.0035	0.0187	11.609	0.0293	0.0711***	11.995	-0.0400**	-0.0400*
		(0.012)	(0.014)		(0.019)	(0.015)		(0.017)	(0.020)
Parish FE		✓	✓		✓	✓		✓	✓
QOB×YOB FE		✓	✓		✓	✓		✓	✓
SES Effects		✓	✓		✓	✓		✓	✓
School Reforms		✓	✓		✓	✓		✓	✓
Parish Trends			✓			✓			✓

Note: *** p <0,01; ** p <0,05; * p <0,1, Standard errors are clustered at the parish level. Top income refers to belonging to the top 20 per cent of the earnings distribution. Covariates which are included in all specifications are a dummy indicating twin births, a dummy for being female, dummies capturing old (>35 years) and young (<20) mothers, a dummy for married women, a dummy indicating a hospital birth and the treatment effect of the maternal intervention. *Mean* refers to the mean value of the outcome variable before the intervention took place. *QOB× YOB effects* include quarter-of-birth dummies for each of the 20 quarters. *Parish FE* are fixed effects for the parish the individual lived in at the time of the birth. *SES effects* are fixed effects for the professional group of the parental household head. *School reforms* refers to the extension of compulsory schooling and length of school year reforms and *Parish trends* are parish specific linear trends.

Pension income to address measurement error in income. Since income is measured at one time, in 1970, when the sample cohorts are 37-39 years old, it may be sensitive to lifecycle variation in labour supply, important for women on account of fertility. We therefore investigated pension income at age 71 as an alternative measure of income. We identify increases in pension income for women of 7%, and no increase for men (lower row, Table 6), ratifying the earnings results. [32] A potential concern with the use of the pension variable is that a widow pension was

[32]For men, we estimate a decline in pension income of 4%. Since we saw no decline in earnings for men at age 37-39, and since only 63% survive to the age of 75, this may reflect endogenous survival selection, the marginal surviving individual being negatively selected post-intervention (see Bhalotra et al. (2017)). To investigate the role of survival selection, we re-estimated programme effects on 1970 income for subsamples of individuals surviving until age 40, 50, 60, 70 and 75 respectively (Appendix Table H1). We see no selection among females until age 75, when there appears to be some positive selection. In contrast, among men, there appears to be negative selection from age 60 onwards as the earnings estimates become progressively lower the older the age group.



(a) Men (b) Women

Note: Covariates which are included are a dummy indicating twin births, dummies capturing old (>35 years) and young (<20) mothers, a dummy for married women, Parish FE and QOB×YOB FE. 90% Confidence Intervals included. Including zero income as an alternative to the log income transformation we do in the main section in Table 6 (see Appendix B).

Figure 7. Quantile regression of income by gender.

available to the sample cohorts, and this could create a wedge between women's earnings and their pensions. However the results are robust to controlling for an indicator for whether the individual was in receipt of a widow pension (see Appendix Table H2).

Internal rate of return. The intervention cost approximately SEK 41,400 (USD 139,000 in current prices). Personnel costs (salaries for physicians and nurses) accounted for 50 per cent of total costs. The cost per treated child was about USD 39 (in current prices) and per consultation USD 5.7 (in current prices). These costs are low relative to the benefits we identify. We obtained average earnings by age, sex and year for 1950-2002 and used them to calculate the net present value of earnings and the internal rates of return for the infant intervention. Details are in Appendix I. The internal rate of return on the funds spent by the national government was about 0.18.

4.2 Intermediate Outcomes- Illuminating Mechanisms

In this section we investigate impacts of the infant health intervention on school-age and labour market outcomes that potentially mediate the observed impacts on test scores and income.

4.2.1 Sickness Absence in Primary School

There are two channels through which infant health may have had the noted impacts on school performance at age 10. First, infant health may predict school-age health, creating a contemporaneous effect from healthy children missing school less often or concentrating better when at school. The second channel operates through brain development and runs directly from infant health to later life cognitive performance (see references in the Introduction). We investigated intervention effects on sickness absence, as a marker of school-age health, with a view to discriminating between the two channels.

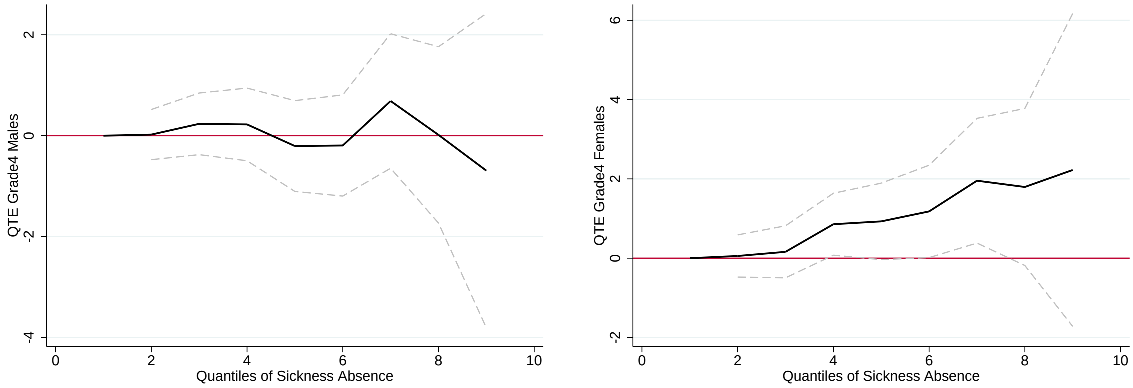
Focusing on Grade 4, where we saw intervention effects on performance, Table 7 and Figure 8 show that the intervention reduced sickness absence for boys by about 0.8% (20%), and it is possible this contributed to their higher GPAs. However, we see an unexpected increase in sickness absence for girls.[33] Given that we find increases in secondary school and earnings for girls and not boys, this undermines the relevance of the pathway involving morbidity in the school years in favour of the argument that the intervention improved neurological development in infancy. We underline that this interpretation is only suggestive.

Table 7. Sickness Absence (Fraction of School Year) – Primary School

	Boys & Girls ($N = 13,138$)			Girls ($N = 6,487$)			Boys ($N = 6,651$)		
	Mean	(1)	(2)	Mean	(3)	(4)	Mean	(5)	(6)
Sickness Absence	0.045	-0.0014	-0.0002	0.050	0.0071	0.0100*	0.040	-0.0080	-0.0083*
		(0.002)	(0.002)		(0.004)	(0.006)		(0.005)	(0.004)
Parish FE		✓	✓		✓	✓		✓	✓
QOB×YOB FE		✓	✓		✓	✓		✓	✓
School FE		✓	✓		✓	✓		✓	✓
SES Effects		✓	✓		✓	✓		✓	✓
Length of Schoolyear		✓	✓		✓	✓		✓	✓
Schoolform		✓	✓		✓	✓		✓	✓
Parish Trends			✓			✓			✓

Note: *** p <0,01; ** p <0,05; * p <0,1, Standard errors are clustered at the parish-grade level. Covariates which are included in all specifications are a dummy indicating twin births, a dummy for being female, dummies capturing old (>35 years) and young (<20) mothers, a dummy for married women, a dummy indicating a hospital birth and the treatment effect of the maternal intervention. *Pre-mean* refers to the mean value of the outcome variable before the intervention took place. *QOB× YOB effects* include quarter-of-birth dummies for each of the 20 quarters. *Parish FE* are fixed effects for the parish the individual lived in at the time of the birth. *SES effects* are fixed effects for the professional group of the household head. *Length of schoolyear* are fixed effects controlling for the reforms concerning the length of the school year. *Schoolform* are fixed effects controlling for the school form as described in Section 2.2 and *Parish specific linear trends* allows for parish specific time trends.

[33] An improvement in health for boys relative to girls is consistent with the stylized fact of boys being more sensitive to health inputs in infancy, although baseline sickness absence rates are similar for boys and girls at about 5% of school days. The distribution and mean of sickness absence for this 1930s births sample resembles closely that in contemporary research (Aucejo and Romano (2014), Goodman (2014)), see Cattan et al. (2017) for further analysis for our sample.



(a) Boys (b) Girls

Note: Covariates which are included are a dummy indicating twin births, dummies capturing old (>35 years) and young (<20) mothers, a dummy for married women, Parish FE and QOB×YOB FE. 90% Confidence Intervals included.

Figure 8. Quantile Regression: Sickness Absence in Grade 4

4.2.2 Employment

Women exposed to the intervention for a year exhibited an increase in the propensity to work full-time of 7.6 percentage points, or 20.5%, and no change in the propensity to work part-time (Table 8). More women joined the labour force, this is explicit in the next section.[34] There are no significant impacts on employment for men, 92.5% of whom worked full-time.

Relative demand for women. In the years when our sample cohorts were making the relevant decisions, there was a substantial expansion of the welfare state and a sharp increase in the share of working married women (Schånberg, 1993). As discussed in the Introduction, we posit that these phenomena are related, as the growing welfare state created more jobs for women than for men, as nurses, teachers and child-care workers. Although our estimates exploit a discontinuity in intervention eligibility conditional upon general trends, we argue it was important that intervention-treated individuals emerging on the market with enhanced skills, faced an expansion of job opportunities. In a similar vein, Coles and Francesconi (2017) argue that expanding job opportunities for women was critical to realisation of the impacts of the contraceptive pill on women's outcomes in America. To investigate the role of opportunities, albeit indirectly, we examined the sectors that women responding to the intervention joined, and linked this to historical information on sectoral growth trends. We discuss these results in the next section.

[34] Part-time refers to 20-35 hours per week and full-time work to more than 35 hours. Both are thought to be underestimated in the 1970 population and household census (cf. Population and Housing Census 1970, 1972b), but this applies to men and women and the under-estimation is unlikely to be correlated with the infant intervention.

Table 8. Employment

	Men & Women (N=20,722)			Women (N=10,256)			Men (N=10,466)		
	Mean	(1)	(2)	Mean	(3)	(4)	Mean	(5)	(6)
Working Parttime	0.145	-0.0201	-0.0147	0.265	-0.0325	-0.0244	0.019	-0.0077	-0.0049
		(0.017)	(0.017)		(0.030)	(0.033)		(0.007)	(0.007)
Working Fulltime	0.640	0.0276	0.0349*	0.370	0.0607*	0.0760**	0.925	-0.0052	-0.0061
		(0.017)	(0.020)		(0.031)	(0.037)		(0.014)	(0.015)
Municipal	0.167	0.0194*	0.0295**	0.238	0.0377*	0.0488**	0.092	0.0012	0.0102
		(0.011)	(0.013)		(0.020)	(0.020)		(0.014)	(0.016)
Governmental	0.081	0.0126	0.0131	0.051	0.0306***	0.0339**	0.111	-0.0053	-0.0077
		(0.013)	(0.014)		(0.012)	(0.014)		(0.019)	(0.019)
Parish FE		✓	✓		✓	✓		✓	✓
QOB×YOB FE		✓	✓		✓	✓		✓	✓
SES Effects		✓	✓		✓	✓		✓	✓
School Reforms		✓	✓		✓	✓		✓	✓
Parish Trends			✓			✓			✓

Note: *** p <0,01; ** p <0,05; * p <0,1, Standard errors are clustered at the parish level. Covariates which are included in all specifications are a dummy indicating twin births, a dummy for being female, dummies capturing old (>35 years) and young (<20) mothers, a dummy for married women, a dummy indicating a hospital birth and the treatment effect of the maternal intervention. *Mean* refers to the mean value of the outcome variable before the intervention took place. *QOB× YOB effects* include quarter-of-birth dummies for each of the 20 quarters. *Parish FE* are fixed effects for the parish the individual lived in at the time of the birth. *SES effects* are fixed effects for the professional group of the parental household head. *School reforms* refers to the extension of compulsory schooling and length of school year reforms and *Parish trends* are parish specific linear trends.

Supply of women- childcare expansion. We have so far emphasized the growing demand for women as workers in the health, education and childcare sectors. We also investigated the extent to which childcare expansion in particular faciliated the labour supply of women. We leverage a major expansion of state-subsidised childcare that occured from 1963, when the exposed cohorts were 30-32 years old, details are in Appendix G. We gathered data on the number of childcare workers per woman of reproductive age in 1960 and 1970 by municipality ($N = 1,037$) as a measure of the intensity of the programme. The number of childcare workers p.c. was close to zero in 1960, and showed significant growth to 1970, at substantially varying rates across municipalities. First, we regressed childcare expansion rates at the local level against the treatment indicator for the infant intervention (treated parish multiplied by eligible birth date), and were able to reject the concern that the childcare expansion was correlated with the original intervention (see Appendix Table G1). We then investigated whether growth in employment and earnings among women exposed to the infant intervention was greater in places where childcare expansion was greater. We found no evidence that it was (see Appendix Table G2).[35] Thus, although the childcare expansion appears not to have facilitated increased labour supply among eligible women, it very likely contributed to the demand for women as childcare workers were predominantly women.

[35]This may be because the expansion only occurred when the women were in their early 30s or it may be, as found in Norway, that the expansion only substituted out informal care (Havnes and Mogstad, 2011).

4.2.3 Occupation

Public Sector Jobs. Using indicators for employment in municipal and central government employment, we find that eligibility for the intervention for a year was associated with an increase in the probability that women work in municipal jobs of 4.9 percentage points, or 20.5% relative to the baseline of about 24% and an increase in the probability of working in central governmental jobs of 3.4 percentage points, or 66.5% (Table 8). Adding up across both categories, it appears that more or less all of the additional employment of women was in the public sector. This lines up with the rapid growth of the Swedish welfare state From the mid-20$^{\text{th}}$ century absorbing women (Stanfors, 2003; Datta Gupta et al., 2006; Sundin and Willner, 2007). Figure 9 shows how female employment rapidly increased from about 800,000 employed women in 1950 to about 1,200,000 in 1970, while male employment stayed fairly constant over time. Figure 10 illustrates the trend in women working in selected public sector jobs 1950–1975.

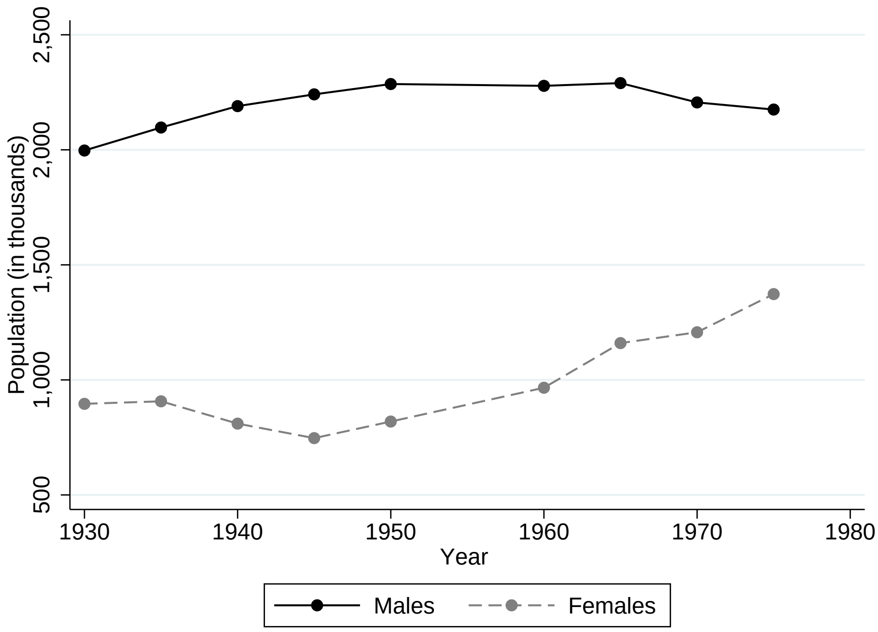


Figure 9. Working population by gender *Source: Statistiska Centralbyrån (2009)*

Occupation. So as to more clearly depict the destinations of women, and to make explcit the skill-content of their tasks, we examined treatment effects on occupation. We find that increases in women's employment were concentrated in high-skilled sectors (Table 9). Women exposed to the intervention for a year were 5.0 percentage points (29.4%) more likely to work as managers and professionals and 4.4 percentage points (35.5%) more likely to work in accounting, banking and administration. In contrast, we see a reduction in the share of men in the professional-management category and an increase in the share of men in sales. Table 9 reports mean earnings

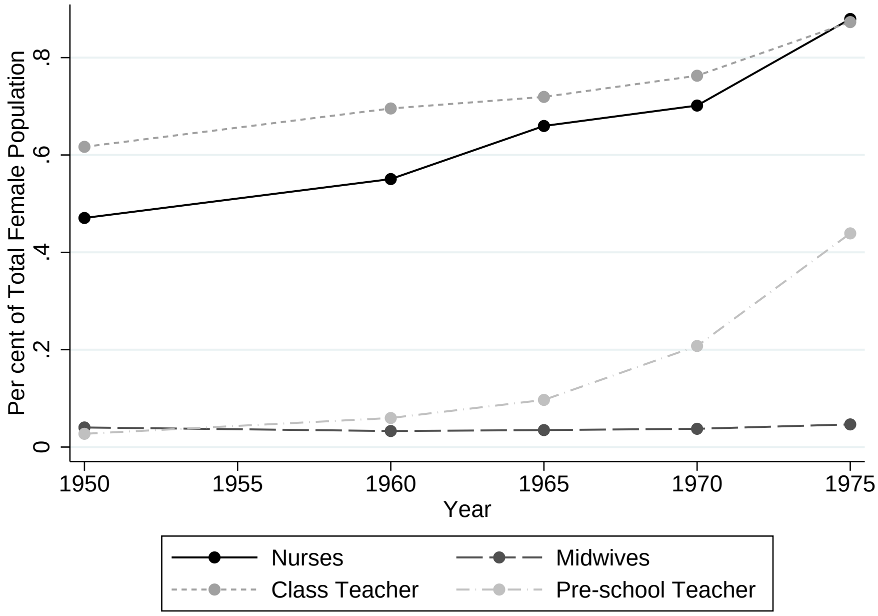


Figure 10. Females working in public sector jobs *Source: Statistiska Centralbyrån (2009).*

by occupation, which confirm that the highest paying occupation was professional-management [36], so these findings line up with the earnings results. Consistent with the employment estimates, these results show a reduction in the *out of the labour force* group for women but not for men. Disaggregating the occupational categories that attracted women further, we find that the largest increase in this category comes from women working in the health sector, for instance as midwifes or nurses (results in Appendix Table J3).

We documented that the infant intervention led to higher school test scores and higher educational attainment for women and we want to link skill accumulation in these intermediate years to the final occupation and earnings outcomes. So in Table 10 we provide descriptive statistics showing, for each occupation, its share and its skill content. We report several indicators of skill content, including share of workers with secondary schooling, average GPA and the average task content classified as routine vs non-routine cognitive vs non-cognitive following Autor et al. (2003). The two highest-ranked occupational groups ('Managers & Professionals' and 'Accounting, administrative') are high-skilled by all three criteria.

[36]Mining shows a higher return for women though not for men. We disregard this aberration as 0.001% of women are in mining.

Table 9. Occupational Sorting

	Men & Women (N=20,920)				Women (N=10,301)				Men (N=10,619)			
	Mean				Mean				Mean			
	Outc.	*Earn.*	(1)	(2)	*Outc.*	*Earn.*	(3)	(4)	*Outc.*	*Earn.*	(5)	(6)
A. Managers, Professionals	0.200	35,473	0.0096	0.0057	0.176	23,909	0.0427**	0.0495***	0.224	44,196	-0.0229	-0.0373**
			(0.014)	(0.011)			(0.019)	(0.019)			(0.021)	(0.019)
B. Accounting, Admin.	0.081	22,408	0.0121	0.0114	0.124	18,825	0.0388	0.0443*	0.036	32,997	-0.0141	-0.0210
			(0.010)	(0.010)			(0.027)	(0.025)			(0.016)	(0.017)
C. Sales	0.083	23,504	-0.0148	-0.0016	0.083	13,063	-0.0245	-0.0226	0.083	33,742	-0.0052	0.0191*
			(0.014)	(0.009)			(0.018)	(0.017)			(0.014)	(0.011)
D. Agricultural	0.059	18,082	0.0090	0.0078	0.026	3,260	0.0099	0.0070	0.093	21,976	0.0081	0.0085
			(0.007)	(0.008)			(0.007)	(0.007)			(0.012)	(0.014)
E. Mining	0.018	29,146	0.0027	0.0013	0.001	24,678	0.0007	0.0003	0.036	29,266	0.0047	0.0024
			(0.004)	(0.005)			(0.001)	(0.001)			(0.008)	(0.009)
F. Transport, Comm.	0.055	25,173	-0.0041	0.0040	0.031	17,346	-0.0081	-0.0062	0.079	27,522	-0.0002	0.0141
			(0.010)	(0.010)			(0.012)	(0.011)			(0.013)	(0.015)
G. Crafts	0.194	25,075	-0.0169	-0.0224*	0.006	31,335	-0.0206	-0.0161	0.335	26,632	-0.0131	-0.0286
			(0.012)	(0.013)			(0.019)	(0.018)			(0.020)	(0.021)
H. Service	0.086	16,283	0.0098	0.0104	0.130	11,288	-0.0087	-0.0033	0.041	29,953	0.0278	0.0238
			(0.012)	(0.015)			(0.015)	(0.016)			(0.019)	(0.020)
I. Out of LF	0.224	3,390	-0.0074	-0.0166	0.370	2,282	-0.0301	-0.0528**	0.072	9,665	0.0149	0.0190
			(0.012)	(0.013)			(0.024)	(0.026)			(0.014)	(0.015)
Parish FE			✓	✓			✓	✓			✓	✓
QOB×YOB FE			✓	✓			✓	✓			✓	✓
SES Effects			✓	✓			✓	✓			✓	✓
School Reforms			✓	✓			✓	✓			✓	✓
Parish Trends				✓				✓				✓

Note: *** p <0,01; ** p <0,05; * p <0,1, Standard errors are clustered at the parish level. We provide means of the dependent variables as shares of men and women working in the occupational category at baseline (Outc.) and also mean earnings for each occupation (Earn.). Covariates which are included in all specifications are a dummy indicating twin births, a dummy for being female, dummies capturing old (>35 years) and young (<20) mothers, a dummy for married women, a dummy indicating a hospital birth and the treatment effect of the maternal intervention. *Mean* refers to the mean value of the outcome variable before the intervention took place. *QOB× YOB effects* include quarter-of-birth dummies for each of the 20 quarters. *Parish FE* are fixed effects for the parish the individual lived in at the time of the birth. *SES effects* are fixed effects for the professional group of the parental household head. *School reforms* refers to the extension of compulsory schooling and length of school year reforms and *Parish trends* allows for parish specific time trends.

Table 10. Descriptive Statistics: Skills and Task Content by Occupation.

	Share		**Occupational Tasks**					**Grades**
	Occ. Group	Sec. Educ.	Nonr. Manual	Routine Manual	Nonr. Cogn. Interactive	Routine Cog.	Nonr. Cogn. Analytic	GPA
Panel A: Men and Women								
All	0.76	0.20	1.568	3.889	1.772	4.488	3.488	-0.009
SD	0.42	0.40	1.375	1.087	2.596	3.714	1.950	(0.769)
Managers & Professionals	0.20	0.47	1.400	4.224	3.029	3.555	5.301	0.304
Accounting, Admin.	0.07	0.32	0.114	4.841	0.632	7.798	3.273	0.318
Sales	0.07	0.17	0.595	3.511	2.669	0.945	4.580	0.091
Agricultural	0.06	0.05	2.418	2.935	4.189	2.284	3.006	-0.166
Transport, Comm.	0.06	0.09	2.882	3.257	1.191	2.267	2.162	-0.154
Crafts	0.20	0.02	1.856	4.287	0.425	7.988	2.759	-0.321
Service	0.09	0.08	1.511	2.902	0.990	1.329	1.798	-0.066
Panel B: Men and Women / No Secondary Education								
All	0.76	0.00	1.671	3.840	1.468	4.614	3.177	-0.150
SD	0.43	0.00	1.409	1.041	2.376	3.766	1.810	(0.769)
Managers & Professionals	0.13	0.00	1.459	4.329	2.337	3.741	5.052	0.031
Accounting, Admin.	0.06	0.00	0.124	4.879	0.629	7.924	3.303	0.168
Sales	0.07	0.00	0.611	3.532	2.489	0.894	4.502	0.025
Agricultural	0.08	0.00	2.419	2.933	4.143	2.279	2.970	-0.202
Transport, Comm.	0.06	0.00	3.004	3.160	1.103	2.016	2.084	-0.225
Crafts	0.24	0.00	1.859	4.288	0.425	7.985	2.761	-0.339
Service	0.10	0.00	1.495	2.902	0.972	1.335	1.777	-0.117

Note: Notes: Descriptive Statistics for Tasks. **Columns:** (2) Share in Occ. Group 1970 (3) Share with Secondary Education Within Occupational Group (4)-(8) Average Tasks for Occupational Group (9) GPA in Primary School. *Source:* Linked 1970 Census. Own calculations. Occupational Tasks based on Autor et al. (2003).

5 Mediators

Identifying mediators is a central challenge in longitudinal studies of early life interventions (Heckman et al., 2013), requiring either two sources of exogenous variation or strong assumptions regarding the relationship between treatment, mediators and main outcomes. For this reason, it has been customary to report the effects of an intervention on potential mediators alongside effects on the final outcomes of interest, without attempting to weight the contributions of alternative mediators. We provided results using this approach in the preceding section.

In recent years, a number of approaches requiring less restrictive assumptions have been suggested. Identification is typically based on a sequential ignorability condition, which states that the unobserved variables that confound the relationship between the treatment and the mediator are different from those that confound the relationship between the mediator and the outcome, conditional on treatment (cf. Heckman and Pinto, 2015; Huber et al., 2017; Dippel et al., 2017). This independence assumption may be plausible in many settings, but in our case, where most outcomes considered are proxies of human capital, it seems difficult to defend such an assumption. We therefore develop a simple approach that can gauge the *relatedness* of the treatment effect of the intervention over different domains. In essence, we examine whether it is the same sub-populations that contribute to the treatment effects in different domains. Under some relatively plausible assumptions, it is possible to determine the extent to which the treatment effects in different outcome domains overlap.[37]

We complement this analysis with the approach developed by Gelbach (2016), which leverages the omitted variable bias formula to attribute treatment effects across potential mediators. The Gelbach approach does not have the ambition of estimating causal effects, and is essentially agnostic about the causal and temporal ordering of potential mediators. Thus, if the treatment effects on different mediators are strongly correlated, the method may deliver misleading results. For example, if one potential mediator (e.g. high-ranking occupation) is a direct consequence of a mediator that was operative at an earlier stage of the life course (e.g. secondary schooling completion) but more strongly correlated with the main outcome (e.g. earnings), then the Gelbach approach may attribute the treatment effect to the later rather than the earlier life course variable. We attempt to (partially) address this pitfall by using insights from our analysis of correlated effects to formulate a specification for the Gelbach (2016) decomposition.

[37] A similar approach has been used in Deuchert et al. (2016), but their approach requires observing the value of the mediator for treated individuals before treatment, and identification is based on this mediator having no effect on the outcome in the pre-treatment period. Thus, their approach cannot be applied to our research design.

5.1 Attribution of Effects

In Appendix F, we show that the estimated average treatment effect on an interaction between two binary outcomes (i.e. $Y = W \cdot Z$), denoted τ_Y, carries information on how strongly the treatment effects within domains defined by the two binary outcomes W and Z relate. First, we may compare τ_Y to the benchmark value τ_Y^{uc} that it would take on if the treatment effects in the two domains were completely unrelated at the individual level:

$$\tau_Y^{uc} = \tau_W \tau_Z + \tau_W \Pr\left(Z^0 = 1\right) + \tau_Z \Pr\left(W^0 = 1\right), \tag{3}$$

where τ_W and τ_Z are the average treatment effects on the two outcomes W and Z and $\Pr\left(Z^0 = 1\right)$ is the (estimable) counterfactual probability of observing $Z = 0$ in the treatment group in the absence of treatment, and $\Pr\left(W^0 = 1\right)$ is analogously defined.

Table 11 shows the relatedness of the treatment effects of the intervention for a number of outcomes that exhibit significant results for women. The first two columns present the estimated treatment effect on the two outcomes mentioned in the leftmost column. For example, the first row shows that exposure to the intervention is associated with an increase in the probability of scoring a high GPA in primary school (grade 4 top 20%) of 10.55 percentage points, and an increase in the probability of secondary schooling of 5.2 percentage points.[38] The third column presents τ_Y^{uc} which is the benchmark value of τ_Y, the effect on the interacted outcome (top GPA *and* secondary schooling), which would be obtained if the treatment effects were uncorrelated. In this particular example, this is 3.4 percentage points. However, the unrestricted treatment effect for this joint outcome, presented in column (4), is almost twice that number, indicating that the treatment effects on the two outcomes are strongly correlated. The three rightmost columns present the estimated correlation coefficient between the two treatment effects.[39]

Table 11 exhibits some striking patterns. First, the estimated value of τ_Y is always well above the benchmark value τ_Y^{uc}, typically twice as large, suggesting that the treatment effects are strongly correlated for all pairs of outcomes (the correlation coefficient is always greater than 0.5 for the maintained assumption on compliers). Treatment effects on earnings are highly correlated with treatment effects on each of high-ranking occupation (0.60), secondary schooling

[38] These are essentially the results from the previous section. They are slightly different because slightly different samples are occasioned now by the requirement that both outcomes are observed for a given individual.

[39] The baseline estimate is based on the assumption that individuals who are compliers for only one of the outcomes are proportionately drawn from the populations of never-takers and always-takers in the other variable. As a sensitivity check we present estimates in square brackets that are obtained with variations in this assumption, allowing that the compliers for one outcome who are never-takers for the other outcome are either strongly under-represented or strongly over-represented; see Appendix F for details.

(0.58) and top GPA (0.54).[40]

Table 11. Correlated Treatment Effects – Women

Outcome 1	Outcome 2	(1) τ_1	(2) τ_2	(3) τ_Y^{uc}	(4) τ_Y	(5)	(6) corr (τ_{1i}, τ_{2i})	(7)
Top GPA (Primary)								
	Secondary	0.1055*	0.0519*	0.0337	0.0664***		0.7738	
		(0.062)	(0.027)		(0.024)	[0.5332	–	0.8426]
	High Occ	0.1044*	0.0631	0.0485	0.0856**		0.9848	
		(0.063)	(0.056)		(0.039)	[0.9524	–	0.996]
	Top Income	0.1044*	0.0837*	0.0465	0.0704*		0.5420	
		(0.063)	(0.050)		(0.041)	[0.1697	–	0.6484]
Secondary Schooling								
	High Occ	0.0396**	0.0815**	0.0276	0.0458***		0.6121	
		(0.017)	(0.038)		(0.014)	[0.2426	–	0.7177]
	Top Income	0.0396**	0.0649**	0.0212	0.0392***		0.5825	
		(0.017)	(0.033)		(0.013)	[0.2213	–	0.6857]
High Occupation								
	Top Income	0.0817**	0.0650**	0.0376	0.0568**		0.6005	
		(0.038)	(0.033)		(0.024)	[0.2748	–	0.6936]

Note: τ_Y^{uc}: benchmark value, uncorrelated effects (see Appendix F for a derivation); τ_1: treatment effect outcome 1; τ_2: treatment effect outcome 2; τ_Y: joint treatment effect for interacted outcome 1×2; corr (τ_{1i}, τ_{2i}): Correlation coefficient between treatment effects (Bounds for alternative assumptions in square brackets; see Appendix F for a derivation).

The preceding results suggest that a plausible sequence of events leads from better primary school performance to secondary school completion and hence better occupations and higher earnings. However, the intervention had larger impacts on primary school scores, occupation and earnings than it did on secondary school completion.[41] This suggests there may be an alternative sequence leading directly from test scores to higher earnings, independent of secondary schooling. So as to discriminate between the two paths, in the next section we estimate a relatively flexible specification that introduces interactions with secondary schooling.

5.2 Gelbach Mediation Analysis

We used the Gelbach (2016) approach to estimate the relative contribution of endogenous outcomes at different stages of the lifecourse to earnings in adulthood. Denoting by Y a $N \times 1$ vector representing log earnings and by T the $N \times 1$ a vector of treatment assignment, we may compare results from two specifications; one where all potential mediators Z are included as

[40] The strongest correlation in effects is found between top GPA in primary school and high-ranking occupations (defined as managers and professionals, and accounting and administration), the correlation coefficients are greater than 0.98 and robust to different assumptions regarding the distribution of compliers in the population. The next largest correlation between treatment effects for top GPA and secondary schooling, at 0.77.

[41] We showed earlier that the intervention had the following impacts: probability of a top GPA increases by about 10 percentage points and the probability of earning a top income increases by 7-8 percentage points, but secondary school completion increases by only 4 percentage points.

covariates, and a base specification which only includes the base covariates and fixed effects X:

$$Y = T\tau + X\lambda + \epsilon \tag{4}$$

$$Y = T\tau + Z\beta + X\lambda + \upsilon \tag{5}$$

Let $\hat{\tau}_{base}$ denote the estimate of τ based on specification (4), and $\hat{\tau}_{full}$ denote the estimate of τ based on specification (5). As shown by Gelbach (2016), their difference $\hat{\delta} = \hat{\tau}_{base} - \hat{\tau}_{full}$ represents an estimate of how much of the estimated effect can be attributed to the mediating variables Z. This decomposition of the effect does not have a causal interpretation since the exogeneity assumption $\mathbb{E}(\upsilon \mid T, Z, X) = 0$ may be violated even if the base specification (4) is identified. Nevertheless, the decomposition gives an indication of the quantitative importance of different potential mediators and their respective contributions to the overall treatment effect τ_{base}. The contribution of variable k can be quantified as $\hat{\delta}_k = \hat{\Gamma}_k \hat{\beta}_k$; where $\hat{\Gamma}_k$ represents the effect of the intervention on mediator k, and $\hat{\beta}_k$ is the estimate for this variable in specification (5).

Table 12 presents results for women.[42] In the top panel, we present estimates of the effects of the infant health intervention on a series of endogenous outcomes, and in the lower panel, we show how these can be attributed to potential mediators. The estimates show that intervention-led improvements in the chances of scoring in the top quintile of the primary school test score distribution account for about two-thirds (66.12%) of the increase in secondary school completion. Secondary schooling, in turn, accounts for a third (28.92%) of the increase in the chances of being in a high-ranking occupation. Secondary education seems key since having a top GPA but failing to achieve secondary schooling has a much smaller and imprecisely determined impact on high occupation. Holding a high-ranking occupation and having secondary schooling explains about half (48.39%) of the intervention-led increase in earnings. Slightly less than half of the increase in earnings is unexplained by the specification.[43]

In each case, the stated links in the chain are the strongest links and this appears to be the predominant trajectory. Although we cannot claim that this is a causal chain, the temporal ordering of the outcomes suggests that, for these early cohorts, the health intervention led to increases in cognitive ability and that the role of primary school performance in determining

[42] Results for an alternative specification which does not interact mediators can be found in Appendix Table J4 and estimates for men, who exhibited no increase in earnings, are in Appendix Table J5.

[43] We have outlined a sequence of extensive margin responses but earnings may also have responded to intensive margin changes, for example, health and cognition related increases in productivity of individuals who would have followed the secondary school pathway irrespective of the intervention.

Table 12. Gelbach Mediation Females.

	Secondary Schooling			High Occupation			Earnings		
Treatment Effect			0.0484*			0.0605			0.1861**
SE			(0.027)			(0.057)			(0.091)
N			6,105			6,105			6,105
Pre-mean			0.189			0.318			9.036
Unexplained =									
Treatment Effect - $\hat{\delta}$			0.0164			0.0392			0.0854
	$\hat{\Gamma}$	$\hat{\beta}$	$\hat{\delta} = \hat{\Gamma} \times \hat{\beta}$	$\hat{\Gamma}$	$\hat{\beta}$	$\hat{\delta} = \hat{\Gamma} \times \hat{\beta}$	$\hat{\Gamma}$	$\hat{\beta}$	$\hat{\delta} = \hat{\Gamma} \times \hat{\beta}$
Top GPA	0.1073*	0.2951***	0.0320*						
	(0.063)	(0.016)	(0.018)						
Secondary Schooling				0.0484*	0.3652***	0.0177*	0.0484*	-0.0118	-0.0006
				(0.027)	(0.036)	(0.011)	(0.027)	(0.099)	(0.005)
Top GPA & Secondary				0.0662***	-0.0030	-0.0002	0.0662***	0.1540***	0.0102**
				(0.024)	(0.039)	(0.003)	(0.024)	(0.034)	(0.005)
Top GPA & No Secondary				0.0422	0.0910***	0.0038	0.0422	0.0016	0.0001
				(0.046)	(0.018)	(0.004)	(0.046)	(0.039)	(0.002)
High Occ & Secondary							0.0596**	1.5088***	0.0899**
							(0.029)	(0.091)	(0.041)
High Occ & No Secondary							0.0009	1.2271***	0.0011
							(0.043)	(0.038)	(0.052)

Note: $\hat{\beta}$ refers to estimates from full model of interest (dependent variable see columns); $\hat{\Gamma}$ refers to estimates from auxiliary models with each possible mediator acting as dependent variable; $\hat{\delta}$ is component of omitted variable bias estimated to be due to each variable (see Gelbach, 2016).

secondary school and the role of secondary school in determining occupational status were probably important pathways to intervention effects on earnings. These results cohere with our finding that the improvements in all of the post-primary outcomes (namely, secondary schooling, occupation and earnings) were unique to women.

6 Robustness Checks

Some robustness checks are incorporated in the main results, for example, we investigated if measurement error in women's earnings was possibly a reason for the gender difference in earnings estimates by using pension income. Here we present further checks.

Treatment indicator. As discussed in the Empirical Strategy section, we investigate alternative (binary) treatment indicators. The pattern of results is in general robust to these variations, and we learn that early exposure, at 0-3 months, is most effective in modifying outcomes (see Appendix Table H4).

Anchoring of grading scale. We investigated sensitivity of the results for academic achievement to the grading scale by anchoring the scale to log income in 1970 (Bond and Lang, 2013; Cunha and Heckman, 2008). We regress income Y for individual i in 1970 on each mark N of the 7-point grading scale for each subject s in each grade g:

$$\ln Y_{isg} = \alpha + \beta^N \text{Mark}_{isg}^N + \varepsilon$$

This gives mean log income for each of the seven steps on the scale. Results are in Appendix Table H5; mark 3 is the reference group. The correlations in fourth grade imply that a switch of test scores from 1 to 6 points is associated with an earnings gain of 95%. Appendix Table H6 shows regression results using income. The estimates are similar to those using the grade scale. Anchoring with years of education instead of log income generates similar results.

Selective survival. To investigate the role of survival selection, we estimate intervention effects on earnings for sub-samples of individuals surviving until age 40, 50, 60, 70 and 75 respectively (see Appendix Table H1). Our main findings are not significantly different across the samples.

Pre-trends test. To investigate whether the outcomes of interest followed similar trends in the treatment and control regions before the intervention, we use the pre-intervention sample to estimate the following equation:

$$y = \beta(trend \times treated) + \gamma treated + \delta trend + \varepsilon.$$

trend is a trend variable based on each $month \times year$ observation in the pre-intervention sample and *treated* is an indicator for treated parishes. A premise of our strategy is that β is 0. Results for men and women are in Appendix Tables H7 and H8 for primary school and labour market outcomes respectively. In general we cannot reject that β is 0. The fact that our results are robust to parish specific time trends also suggests the absence of differential pre-trends.

Attrition. Match rates are reported in Section 3. Match rates of the birth sample with the 1970 census and the tax register are high, but the dominant cause of attrition is death. Since we know from Bhalotra et al. (2017) that the intervention of interest altered child and adult mortality risks, there is a potential concern about differential attrition. The match rate with the school data is lower since, in addition to death, some of the regional archives did not preserve all school catalogues. In Appendix Table H9 we provide an overview of attrition rates for sub-samples identified by treated parish and eligible birth date. Table 13 provides tests for whether attrition is systematically related to treatment. We find that attrition is negatively associated with treatment in the school sample, but that it is uncorrelated with treatment in the later-life samples.

We address this in two ways. First, we ran the analysis of census outcomes on a sample restricted to individuals who also appear in the school sample (see Appendix Table H10). We

Table 13. Attrition in Different Samples

	School Sample			Census 70 Sample			Pension Sample		
Treated Parish	0.0450			0.0014			-0.0167		
	(0.047)			(0.012)			(0.015)		
Duration Eligibility	0.0200	0.0985	0.1048	0.0193**	0.0133	0.0142	0.0146	0.0553	0.0458
	(0.015)	(0.077)	(0.076)	(0.008)	(0.074)	(0.073)	(0.021)	(0.087)	(0.087)
Treated × Duration Eligibility	-0.0387*	-0.0370*	-0.0562**	-0.0022	-0.0025	-0.0100	0.0007	-0.0014	-0.0101
	(0.021)	(0.020)	(0.023)	(0.009)	(0.009)	(0.009)	(0.023)	(0.021)	(0.020)
Baseline	0.339	0.339	0.339	0.140	0.140	0.140	0.336	0.336	0.336
Parish FE		✓	✓		✓	✓		✓	✓
QOB×YOB FE		✓	✓		✓	✓		✓	✓
School FE		✓	✓		✓	✓		✓	✓
SES Effects		✓	✓		✓	✓		✓	✓
Length of Schoolyear		✓	✓						
Schoolform		✓	✓						
Parish Trends			✓			✓			✓

continue to see positive and significant effects of the intervention on women's education and earnings of magnitudes similar to the baseline estimates.[44] Second, we estimated Lee bounds for all main outcomes. See Appendix Tables H11 and H12. The results stand up to accounting for attrition in this way. With the one exception of "Top GPA", for which the lower Lee bound has a p value of 0.11, all of the positive effects reported for women are significant even at the lower Lee bound.

Placebos. For the 1970 census outcomes we implement a placebo test, using a fake intervention ten years after the actual intervention. For this, we generated a sample of children born in treatment and control areas ten years after the infant intervention, in 1940-44, using the 1950 population census and the Swedish Death Index. The only covariate available from these sources is the individual's sex. Generating a fake treatment group using parish and date of birth, we show that the fake intervention had no long run impacts (Table 14).[45]

Randomisation Inference. We conduct a randomisation inference test for the long-term outcomes, also in the spirit of a placebo test. We randomly assign treatment status within each treatment and control parish pair using 5,000 permutations; cf. Karlsson and Pichler, 2015 for a discussion of randomisation inference in difference-in-difference settings. Following MacKinnon and Webb (2016) we present results based on t statistics, as this is superior to

[44]The exercise in Section 5 was also on this restricted sample, where we saw the results hold. We just make this more explicit here.

[45]In order to use the years 1934-1939, we would need to digitize additional years of parish birth records and school exam catalogues. The results for the 1940–1944 cohorts have to be viewed with some caution since before 1947 the parish of birth that was reported refers to the location of the hospital the birth was in and not to the place of registration of the parents (Holmlund, 2008). With a rising share of institutionalised births over time this leads to some misreporting for our placebo test cohorts. We do not face this problem for cohorts born 1930–1934 since the parish records that were digitised within this project report the place of registration of the parents and not the place of the hospital they were born in. To mitigate the problem, we control for hospital births.

Table 14. Placebo 1940-1944 cohorts long-term outcomes.

	Secondary Schooling		**Working Fulltime**		**Working Parttime**	
	(1)	(2)	(3)	(4)	(5)	(6)
DID	-0.0005	-0.0013	-0.0002	-0.0016	0.0003	0.0008
SE	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)
N	19,110	19,110	19,339	19,339	19,339	19,339
Pre-Mean	0.231	0.231	0.603	0.603	0.071	0.071
	log Income 1970		**Municipal**		**Governmental**	
	(1)	(2)	(3)	(4)	(5)	(6)
DID	0.0017	-0.0013	-0.0001	0.0005	-0.0007	-0.0013
SE	(0.004)	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)
N	19,634	19,634	19,339	19,339	19,339	19,339
Pre-Mean	9.391	9.391	0.140	0.140	0.079	0.079
QOB×YOB Effects	✓	✓	✓	✓	✓	✓
Parish Specific Linear Trends		✓		✓		✓

Note: Standard errors clustered at the parish level in parenthesis. DID denotes term for placebo exposure to the infant intervention assuming it was implemented 10 years later in the same areas. Apart from *QOB×YOB effects* which indicate inclusion of quarter-of-birth dummies for each of the 20 quarters and *Parish specific trends* testing for parish specific time trends, the only control variable in all specifications is a dummy for being female.

inference based on coefficients. Figures 11 and 12 plot the distributions of placebo treatment effects by gender and display the actual treatment effect and the corresponding p value. Except for part-time employment for females where the distribution does not look smooth (and for which we concluded earlier that there was no significant intervention effect), the results are similar to the main estimates in Tables 4, 6 and 8.

Figure 11. Randomisation inference long-term outcomes females.

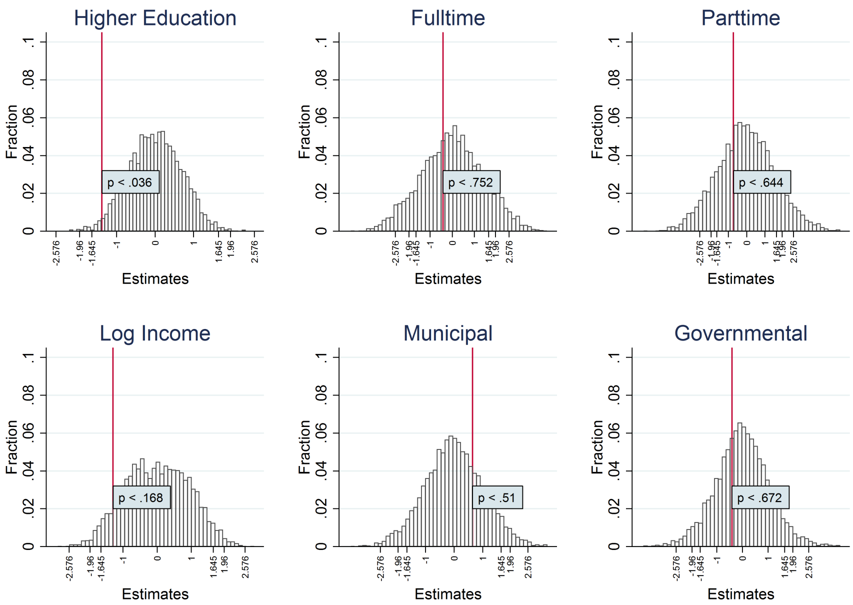


Figure 12. Randomisation inference long-term outcomes males.

7 Conclusion

Using unique longitudinal data in which individual outcomes are observed at different stages of the life course, from birth through to pension age, we identify large impacts of a universal infant care intervention on school and labour market outcomes. A crude estimate of the internal rate of return suggests that the intervention was highly cost-effective, and it was successfully scaled up following the short trial period that we analyse. Our findings are of contemporary relevance given that poor health and nutrition and deficient early childhood care are estimated to be causing about 200 million children under the age of 5 to fail to attain their cognitive potential, and that this has been identified as a key factor in the intergenerational transmission of poverty (Grantham-McGregor et al., 2007).

Intervention effects are highly correlated across outcomes, implying that it is largely the same individuals who drive the various effects. With the caveat that it is only descriptive, our analysis of mediators suggests that cognitive attainment and secondary schooling were important contributors to the increase in adult earnings. The analysis also highlights the importance of institutional capacity (rationing of secondary school places) and demand conditions (demand for women's labour). Entry to secondary school was competitive and this led to more girls and not boys progressing into secondary schooling as a result of the intervention. When these cohorts emerged onto the labour market, the movement of women into high-earning sectors is likely to have been facilitated by substantial growth in women-friendly public sector jobs created by rapid expansion of the welfare state.

References

Acemoglu, D., D. H. Autor, and D. Lyle (2004). Women, war, and wages: The effect of female labor supply on the wage structure at midcentury. *Journal of political Economy 112*(3), 497–551.

Adhvaryu, A. and A. Nyshadham (2016). Endowments at birth and parents' investments in children. *The Economic Journal 126*(593), 781–820.

Aizer, A., J. Currie, P. Simon, and P. Vivier (2018). Do low levels of blood lead reduce children's future test scores? *American Economic Journal: Applied Economics 10*(1), 307–41.

Almond, D. and J. Currie (2011). Killing me softly: The fetal origins hypothesis. *Journal of Economic Perspectives 25*(3), 153–172.

Almond, D., J. Currie, and V. Duque (2017). Childhood circumstances and adult outcomes: Act ii. *Forthcoming Journal of Economic Literature*.

Almond, D., L. Edlund, and M. Palme (2009). Chernobyl's Subclinical Legacy: Prenatal Exposure to Radioactive Fallout and School Outcomes in Sweden. *The Quarterly Journal of Economics 124*(4), 1729–1772.

Almond, D., B. Mazumder, and R. Van Ewijk (2014). In utero ramadan exposure and children's academic performance. *The Economic Journal 125*(589), 1501–1533.

Attanasio, O. P. (2015). The determinants of human capital formation during the early years of life: Theory, measurement, and policies. *Journal of the European Economic Association 13*(6), 949–997.

Attanasio, O. P., C. Fernandez, E. Fitzsimons, S. Grantham-McGregor, C. Meghir, and M. R. Codina (2014). Using the infrastructure of a conditional cash transfer program to deliver a scalable integrated early child development program in colombia: cluster randomized controlled trial. *British Medical Journal 349*(g5785).

Aucejo, E. and T. Romano (2014). Assessing the Effect of School Days and Absences on Test Score Performance. CEP Discussion Papers 1302, Centre for Economic Performance, London School of Economics.

Autor, D. H., F. Levy, and R. J. Murnane (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly journal of economics 118*(4), 1279–1333.

Baird, S., F. H. Ferreira, B. Özler, and M. Woolcock (2014). Conditional, unconditional and everything in between: a systematic review of the effects of cash transfer programmes on schooling outcomes. *Journal of Development Effectiveness 6*(1), 1–43.

Baird, S., J. H. Hicks, M. Kremer, and E. Miguel (2016). Worms at work: Long-run impacts of a child health investment. *The Quarterly Journal of Economics 131*(4), 1637–1680.

Baker, M., J. Gruber, and K. Milligan (2018). The long-run impacts of a universal child care program. *American Economic Journal: Economic Policy Forthcoming.*

Banerjee, A. V., S. Cole, E. Duflo, and L. Linden (2007). Remedying Education: Evidence from Two Randomized Experiments in India. *The Quarterly Journal of Economics 122*(3), 1235–1264.

Bång, J. (2001). The Returns to Education - Using Data Retrieved from the Swedish National Census of 1930. *mimeo, Department of Economics, Uppsala University*.

Bergh, A. (2009). *Den kaptalistiska välfärdsstaten - om den svenska modellens historia och framtid*. Norstedts förlag.

Bernhardtz, L. and Klintfelt, A. (2007). Den typiska adoptivbarnet - en svenskfödd 40-talist. *Välfärd 2*.

Bhalotra, S., M. Fernandez-Sierra, and A. S. Venkataramani (2018). Women's labour force participation and the gender wage gap across the distribution of tasks.

Bhalotra, S., M. Karlsson, and T. Nilsson (2017). Infant health and longevity: Evidence from a historical intervention in sweden. *Journal of the European Economic Association*, jvx028.

Bhalotra, S., A. Venkataramani, and S. Walther (2018). Fertility responses to reductions in mortality: Quasi-experimental evidence from 20th century america.

Bhalotra, S. R. and A. S. Venkataramani (2012). Shadows of the captain of the men of death: Early life health interventions, human capital investments, and institutions. *Available at SSRN*.

Bhalotra, S. R. and A. S. Venkataramani (2013). Cognitive development and infectious disease: Gender differences in investments and outcomes. *IZA Discussion Paper* (7833).

Bharadwaj, P., J. Eberhard, and C. Neilson (2017). Health at birth, parental investments and academic outcomes. *Forthcoming in Journal of Labour Economics*.

Bharadwaj, P., K. V. Løken, and C. Neilson (2013). Early life health interventions and academic achievement. *American Economic Review 103*(5), 1862–1891.

Bitler, M., T. Domina, and H. Hoynes (2016). Experimental Evidence on Distributional Effects of Head Start. Mimeo, UC Davis.

Björklund, A. and C. Kjellström (1994). Avkastningen på utbildning i sverige 1968 till 1991.

Black, S. E., P. J. Devereux, and K. G. Salvanes (2007). From the cradle to the labor market? the effect of birth weight on adult outcomes. *The Quarterly Journal of Economics*, 93–120.

Bobonis, G. J., E. Miguel, and C. Puri-Sharma (2006). Anemia and school participation. *Journal of Human resources 41*(4), 692–721.

Bond, T. N. and K. Lang (2013). The Evolution of the Black-White Test Score Gap in Grades K–3: The Fragility of Results. *The Review of Economics and Statistics 95*(5), 1468–1479.

Bozzoli, C., A. Deaton, and C. Quintana-Domeque (2009). Adult height and childhood disease. *Demography 46*(4), 647–669.

Broman, I. T. (1995). *Perspektiv på förskolans historia.* Studentlitteratur.

Bütikofer, A., K. V. Løken, and K. G. Salvanes (2019). Infant health care and long-term outcomes. *Review of Economics and Statistics* (Forthcoming).

Cattan, S., G. Conti, and C. Farquharson (2019). Workforce quality in early years intervention: Evidence from a large-scale home-visiting program.

Cattan, S., D. A. Kamhöfer, M. Karlsson, and T. Nilsson (2017). The short-and long-term effects of student absence: Evidence from sweden.

Chay, K. C., J. Guryan, and B. Mazumder (2009). Birth cohort and the black-white achievement gap: The roles of access and health soon after birth. (Working paper 15078).

Chetty, R., J. N. Friedman, N. Hilger, E. Saez, D. W. Schanzenbach, and D. Yagan (2011). How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star. *The Quarterly Journal of Economics 126*(4), 1593–1660.

Clarke, D., G. Cortés Méndez, and D. Vergara Sepúlveda (2018). Growing together: Assessing equity and efficiency in an early-life health program in chile.

Coles, M. G. and M. Francesconi (2017). Equilibrium search and the impact of equal opportunities for women. *Journal of Political Economy*.

Cortes, G. M., N. Jaimovich, and H. E. Siu (2018). The" end of men" and rise of women in the high-skilled labor market. Technical report, National Bureau of Economic Research.

Cunha, F. and J. J. Heckman (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *The Journal of Human Resources 43*(4).

Cunha, F., J. J. Heckman, and S. M. Schennach (2010, 05). Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Econometrica 78*(3), 883–931.

Dahr, E. (1945). Lärjungevalet till studielinjer med den nuvarande realskolans mål. *SOU 1945 44*.

Datta Gupta, N., N. Smith, and M. Verner (2006). Child Care and Parental Leave in the Nordic Countries: A Model to Aspire to? IZA Discussion Papers 2014.

De Chaisemartin, C. (2012). Fuzzy differences in differences. Pse working papers, HAL.

Deuchert, E., M. Huber, and M. Schelker (2016). Direct and indirect effects based on difference-in-differences with an application to political preferences following the vietnam draft lottery.

Deverman, B. E. and P. H. Patterson (2009). Cytokines and Cns Development. *Neuron 64*, 61–78.

Dhamija, G. and S. Gitanjali (2019). Impact of early life interventions on later life outcomes: Evidence from india's integrated child development services.

Dippel, C., R. Gold, S. Heblich, and R. Pinto (2017). Instrumental variables and causal mechanisms: Unpacking the effect of trade on workers and voters. Technical report, National Bureau of Economic Research.

Doyle, O., C. P. Harmon, J. J. Heckman, and R. E. Tremblay (2009). Investing in early human development: Timing and economic efficiency. *Economics & Human Biology 7*(1), 1–6.

Duflo, E. and R. Hanna (2005). Monitoring works: Getting teachers to come to school. NBER Working Papers 11880, National Bureau of Economic Research, Inc.

Edvinsson, R. (2016). Historical currency converter. *URL: http://www. historicalstatistics. org/Currencyconverter. html.[Retrieved on Apr 15, 2018].*

Ekstrand, B. (2000). *Småbarnsskolan: vad hände och varför?: en sekellång historia studerad med fokus på förändring av pedagogisk verksamhet från 1833 och framåt.* Britten Ekstrand, Högskolan Kristianstad, S-291 88 Kristianstad.

Elsby, M. W., B. Hobijn, and A. Sahin (2010). The labor market in the great recession. Technical report, National Bureau of Economic Research.

Engle, P. L., M. M. Black, J. R. Behrman, M. Cabral de Mello, P. J. Gertler, L. Kapiriri, R. Martorell, and M. E. Young (2007). Strategies to avoid the loss of developmental potential in more than 200 million children in the developing world. *The Lancet 369*(9557), 229–242.

Eppig, C., C. L. Fincher, and R. Thornhill (2010). Parasite prevalence and the worldwide distribution of cognitive ability. *Proceedings of the Royal Society of London B: Biological Sciences 277*(1701), 3801–3808.

Falk, A. and F. Kosse (2016). Early childhood environment, breastfeeding and the formation of preferences.

Field, E., O. Robles, and M. Torero (2009). Iodine deficiency and schooling attainment in tanzania. *American Economic Journal: Applied Economics 1*(4), 140–69.

Figlio, D., J. Guryan, K. Karbownik, and J. Roth (2014). The effects of poor neonatal health on children's cognitive development. *American Economic Review 104*(12), 3921–3955.

Finch, C. E. and E. M. Crimmins (2004). Inflammatory exposure and historical changes in human life-spans. *Science 305*(5691), 1736–1739.

Firpo, S., N. M. Fortin, and T. Lemieux (2009). Unconditional quantile regressions. *Econometrica 77*(3), 953–973.

Fischer, M., M. Karlsson, and T. Nilsson (2013). Effects of compulsory schooling on mortality: evidence from Sweden. *International journal of environmental research and public health 10*(8), 3596–3618.

Fischer, M., M. Karlsson, T. Nilsson, and N. Schwarz (2017). The Long-Term Effects of Long Terms. Compulsory Schooling Reforms in Sweden.

Flavio and J. Heckman (2007). The technology of skill formation. *American Economic Review 97*(2), 31–47.

Fredriksson, V., S. Marklund, G. Sivgard, and M. Widen (1971). Svenska folkskolans historia, sjätte delen. skolutvecklingen 1942–1962.

Garcia, J. L., J. J. Heckman, and A. L. Ziff (2018). Gender differences in the benefits of an influential early childhood program. *European economic review 109*, 9–22.

Gelbach, J. B. (2016). When do covariates matter? and which ones, and how much? *Journal of Labor Economics 34*(2), 509–543.

Goldin, C. (1988). Marriage bars: Discrimination against married women workers, 1920's to 1950's.

Goodman, J. (2014). Flaking Out: Student Absences and Snow Days as Disruptions of Instructional Time. NBER Working Papers 20221, National Bureau of Economic Research, Inc.

Gorna, R., N. Klingen, K. Senga, A. Soucat, and K. Takemi (2015). Women's, children's, and adolescents' health needs universal health coverage. *The Lancet 386*(10011), 2371–2372.

Grantham-McGregor, S., Y. B. Cheung, S. Cueto, P. Glewwe, L. Richter, and B. Strupp (2007). Developmental potential in the first 5 years for children in developing countries. *Lancet 369*(9555), 60–70.

Hammarlund, K. (1998). Barnet och barnomsorgen. *Bilden av Barnet i ett Socialpolitiskt Projekt, Gothenburg, Dissertations from Department of History, Gothenburg University no 19.*.

Hatje, A.-K. (1999). *Från treklang till triangeldrama: barnträdgården som ett kvinnligt samhällsprojekt under 1880-1940-talen.* Historiska media.

Havnes, T. and M. Mogstad (2011). Money for nothing? universal child care and maternal employment. *Journal of Public Economics 95*(11-12), 1455–1465.

Heckman, J., D. Ohls, R. Pinto, and M. Rosales (2014). A reanalysis of the nurse family partnership program: The memphis randomized control trial. `https://heckman.uchicago.edu/sites/heckman2013.uchicago.edu/files/uploads/CEHD_Launch/2_CEHD-NFP_SLIDES_2014-05-29a_MR_FINAL.pdf`.

Heckman, J., R. Pinto, and P. Savelyev (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review 103*(6), 2052–86.

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science 312*(5782), 1900–1902.

Heckman, J. J. and R. Pinto (2015). Econometric mediation analyses: Identifying the sources of treatment effects from experimentally estimated production technologies with unmeasured and mismeasured inputs. *Econometric reviews 34*(1-2), 6–31.

Heckman, J. J., J. Stixrud, and S. Urzua (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor economics 24*(3), 411–482.

Hjalmarsson, R., H. Holmlund, and M. J. Lindquist (2015). The effect of education on criminal convictions and incarceration: Causal evidence from micro-data. *The Economic Journal 125*(587), 1290–1326.

Hjort, J., M. Sølvsten, and M. Wüst (2017). Universal investment in infants and long-run health: Evidence from denmark's 1937 home visiting program. *American Economic Journal: Applied Economics 9*(4), 78–104.

Holmlund, H. (2008). A Researchers Guide to the Swedish Compulsory School Reform. CEE Discussion Papers 0087, Centre for the Economics of Education, LSE.

Huber, M., M. Lechner, and A. Strittmatter (2017). Direct and indirect effects of training vouchers for the unemployed. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*.

Imbens, G. and J. Woolridge (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature 47*(1), 5–86.

Johansson, M. et al. (2006). *Inkomst och ojämlikhet i Sverige 1951-2002.* Institutet för framtidsstudier Stockholm.

Karlsson, M. (2002). *Perspektiv på familjedaghem.* Ph. D. thesis, Dissertation, Studies in Educational Sciences, Stockholm University, Stockholm: HLS förlag.

Karlsson, M. and S. Pichler (2015). Demographic consequences of HIV. *Journal of Population Economics 28*(4), 1097–1135.

Kyle, G. and G. Herrström (1972). *Två studier i den svenska flickskolans historia.* Föreningen för svensk undervisningshistoria.

Lang, K. (2010). Measurement matters: Perspectives on education policy from an economist and school board member. *Journal of Economic Perspectives 24*(3), 167–82.

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *The Review of Economic Studies 76*(3), 1071–1102.

Leeuwen, v. M., I. Maas, and A. Miles (2002). Hisco: Historical international standard classification of occupations. *Leuven University Press*.

Lindgren, K.-O., S. Oskarsson, and C. T. Dawes (2014). Can political inequalities be educated away? Evidence from a Swedish school reform. Technical report, Working Paper, IFAU-Institute for Evaluation of Labour Market and Education Policy.

Maccini, S. and D. Yang (2009). Under the weather: Health, schooling, and economic consequences of early-life rainfall. *American Economic Review 99*(3), 1006–26.

MacKinnon, J. G. and M. D. Webb (2016). Randomization Inference for Difference-in-Differences with Few Treated Clusters. Working Papers 1355, Queen's University, Department of Economics.

Maluccio, J. A., J. Hoddinott, J. R. Behrman, R. Martorell, A. R. Quisumbing, and A. D. Stein (2009). The impact of improving nutrition during early childhood on education among guatemalan adults. *The Economic Journal 119*(537), 734–763.

Molina, T. (2016). Pollution, ability, and gender-specific investment responses to shocks.

Muralidharan, K. and V. Sundararaman (2011). Teacher performance pay: Experimental evidence from india. *Journal of Political Economy 119*(1), 39 – 77.

Paulsson, E. (1946). Om folkskoleväsendets tillstånd och utveckling i sverige under 1920- och 1930-talen. *Jönköping: Länstryckeriaktiebolaget*.

Pettersson-Lidbom, P. (2015). Midwives and Maternal Mortality: Evidence from a Midwifery Policy Experiment in Sweden in the 19th Century.

Pitt, M. M., M. R. Rosenzweig, and M. N. Hassan (2012). Human capital investment and the gender division of labor in a brawn-based economy. *The American Economic Review 102*(7), 3531–3560.

Population and Housing Census 1970 (1972a). Part 1. population in communes and parishes, etc. *National Bureau of Statistics*.

Population and Housing Census 1970 (1972b). Part 10. industry, occupation and education in the whole county etc. *National Bureau of Statistics*.

Saaritsa, S. and A. Kaihovaara (2016). Good for girls or bad for boys? schooling, social inequality and intrahousehold allocation in early twentieth century finland. *Cliometrica 10*(1), 55–98.

Schånberg, I. (1993). *Den kvinnliga utbildningsexpansionen 1916-1950: realskolestadiet*, Volume 27.

Sievertsen, H. H. and M. Wüst (2017). Discharge on the day of birth, parental response and health and schooling outcomes. *Journal of health economics 55*, 121–138.

SOU (1929). Betänkande angående moderskapsskydd. *Statens offentliga utredningar, Stockholm 1929:28.*

SOU (1935). Kungl. medicinalstyrelsens utlåtande och förslag angående förebyggande mödra och barnavård. *Statens offentliga utredningar 1935:19.*

SOU (1942). Betänkande med utredning och förslag angående betygsättningen i folkskolan, angivet av inom ecklesiastikdepartementet tillkalade sakkunniga. *Statens offentliga utredningar 1942:11.*

SOU (1944). 1940 års skolutrednings betänkanden och utredningar. *Statens offentliga utredningar 1944:21.*

SOU (1972). Statens offentliga utredning. förskolan, del 2. betänkande av barnstugeutredningen. sou 1972:27.

Stanfors, M. (2003). *Education, labor force participation and changing fertility patterns : a study of women and socioeconomic change in twentieth century Sweden*, Volume 22. Lund studies in economic history.

Statistiska Centralbyrån (2009). Förvärvsarbetande: Folk- och bostadsräkningarna 1910-1985. *National Bureau of Statistics*.

Stenhoff, G. (1931). Huru nedbringa dödligheten bland de späda barnen? *Tidskrift för Barnavård och Ungdomsskydd 6*, 283–288.

Stenhoff, G. (1934). Försöksverksamhet beträffande för- och eftervård vid barnsbörd. *Tidskrift för Barnavård och Ungdomsskydd 3*, 99–101.

Sundén, A. (2006). The swedish experience with pension reform. *Oxford Review of Economic Policy 22*(1), 133–148.

Sundin, J. and S. Willner (2007). *Social change and health in Sweden. 250 years of politics and practice.* Swedish National Institute of Public Health R 2007:21; Alfa Print AB, Solna 2007.

Swedish Government, S. (1931). Government proposition 1931:306. *Stockholm*.

UNESCO (2014). *Teaching and learning: Achieving quality for all*, Volume 2013/4. UNESCO Global Monitoring Report.

Wallgren, A. (1936). Den förebyggande spädbarnsvårdens organisation. *Tidskrift för Barnavård och Ungdomsskydd* (5), 163–67.

Wallin, H. and H. Grimlund (1933). års förnyade läroverksstadga: med förklaringar och hänvisningar: jämte timplaner och undervisningsplan mm rörande allmänna läroverken.

Westberg, J. (2015). Förskolans historia. in. *Larsson & J. Westberg (Red.), Utbildningshistoria: En introduktion*.

World Bank (2015). *Mind, Society and Behaviour. World Development Report 2015.* World Bank Group.

Yi, J., J. J. Heckman, J. Zhang, and G. Conti (2015). Early health shocks, intra-household resource allocation and child outcomes. *The Economic Journal 125*(588), F347–F371.

For Online Publication

A Appendix: Swedish Grading System

The grading scale used throughout the period was introduced in 1897, and was applicable to all subjects but not to behavioural marks (these had a shorter scale and much higher concentration in the highest marks). Officially marks were given on a seven-point grading scale which ranged from A (passed with great distinction) to C (failed). Teachers were also allowed to use + and - signs to express the strength or weakness of a mark. A complete list of applied marks and their meaning can be seen in Table A1. At the outset, there was some heterogeneity in how student performance was evaluated, but our investigation period falls into a period of constantly increasing comparability between schools and teachers in their marking of pupil performance.

A pass mark, i.e. at least a B, was required in theoretical subjects to proceed to the next grade.[46] There was, however, some local variation in how this rule was enforced in practice: some districts required a pass mark in all theoretical subjects; some allowed for a maximum of two fails, provided these two are not Swedish and math. Other districts allowed for very high marks in some subjects to offset low marks in other subjects.

Since from 1939 onwards, admission to secondary school was based on marks from primary school, a Royal Commission emphasised that the marking procedure should be improved and standardised much more. Therefore, guidelines for marking were prepared which were published in 1940 and became official starting with the school year 1940/41. These provided general guidelines for the marks and gave further information on individual subjects. It was stated that marks should be defined in a relative sense, meaning that *Ba* is defined as the normal mark which should encompass the middle third of a pupil's cohort. Consequently, one third of the other pupils should fall below this mark and the other third should be above. Only in really exceptional cases pupils obtained the extreme marks C or A. According to the commission, less than one percent of the pupils could be expected to have the knowledge corresponding to the top mark A, which should testify exceptional talent. As discussed in SOU (1942) the recommendation was to also have *Ba* as the normal mark in grade 1 and grade 2, but the authorities explicitly recommended teachers to be restrictive in the use of any high or low marks for children in these grades.

[46] There are only very few statistics on how common grade retention was at that time, but a survey in 1940 from the second largest city of Sweden Gothenburg suggests that about 3% of all pupils had to repeat a grade (Paulsson, 1946)

Table A1. 1897 grading scale.

Mark	Name	English
A	*Berömlig*	Passed with great distinction
a	*Med utmärkt beröm godkänd*	Passed with distinction
AB	*Med beröm godkänd*	Passed with great credit
Ba	*Icke utan beröm godkänd*	Passed with credit
B	*Godkänd*	Passed
BC	*Icke fullt godkänd*	Not entirely passable
C	*Underkänd*	Fail

Note: Official Swedish grading scale from 1897 as described in Section 2.2 and their English interpretation.

B Appendix: Variable Definitions

B.1 Information from Parish Records

Female Dummy variable taking on the value one for female births.

Twin Dummy variable taking on value one for (mono- and dizygotic) twins.

Wedlock Dummy variable taking on value one for children born to married mothers.

Mother<20 Dummy variable taking on value one for mothers younger than 20 years at time of birth.

Mother>35 Dummy variable taking on value one for mothers older than 35 years at time of birth.

Hospital birth Dummy variable taking on value one for child being born in hospital.

Treated Dummy variable taking on value one for children born in treated areas.

TreatmentI Dummy variable indicating eligibility for infant care intervention during at least the first three months in life.

DurationI Variable indicating eligibility for infant care intervention in years.

TreatmentM Dummy variable indicating eligibility for maternal care intervention during at least the first three months in life.

DurationM Variable indicating eligibility for maternal care intervention in years.

SES Classification of head of household profession according to HISCO 9-point scale (Leeuwen et al., 2002).

B.2 Variables from Exam Catalogues:

Share Sickn. Abs. Share of school days spend absent due to sickness in grade 1 or 4.

Writing Mark for "writing" in grade 1 or 4.

Reading/Speaking Mark for "reading and speaking" in grade 1 or 4.

Math Mark for "math" in grade 1 or 4.

Religion Mark for "religion" in grade 1 or 4.

GPA Grade point average of subjects "math", "reading and speaking" and "writing" in grade 1 or 4.

B.3 Variables from 1970 Population and Household Census:

Only Primary Dummy variable taking on value one for someone having only primary education.

Dropout Secondary Dummy variable taking on value one for someone who attended but did not finish secondary school.

Secondary Schooling Dummy variable taking on value one for someone having higher education than *Folkskola*.

Working Fulltime Dummy variable taking on value one for someone working at least 35 hours per week.

Working Parttime Dummy variable taking on value one for someone working at least 20 but not more than 34 hours per week.

log Income Logarithmised taxable labour earnings. Imputed an income based on qualification and hours worked for those having zero income and made a log+1 transformation for remaining zero incomes.

Top Income 1970 Dummy variable taking on value one for someone at the upper 20% of the earnings distribution.

Municipal (public) Employment Dummy variable taking on value one for someone working in the municipal (public) sector. Lower (local) level of government.

Governmental (public) Employment Dummy variable taking on value one for someone working in the governmental (public) sector. Higher (state) level of government.

Scientific, Medical, Technical Dummy variable taking on value one for someone working in the scientific, medical or technical branch.

Admin. Dummy variable taking on value one for someone working in the administrative branch.

Accounting, Admin. Dummy variable taking on value one for someone working in the accounting branch.

Sales Dummy variable taking on value one for someone working in the sales branch.

Agricultural Dummy variable taking on value one for someone working in the agricultural or fishing branch.

Mining Dummy variable taking on value one for someone working in the mining branch.

Transport, Communication Dummy variable taking on value one for someone working in the transport or communication branch.

Crafts Dummy variable taking on value one for someone working in the crafts branch.

Service Dummy variable taking on value one for someone working in the service branch.

Out of the Labour Force Dummy variable taking on value one for someone being out of the labour force or having a non-identified job.

C Appendix: Matching Procedure

This section provides more details on the matching procedure discussed in Section 3.

The Mahalanobis distance metric is defined as

$$\mathcal{J}_M(i) = \arg\min_j \sqrt{(X_i - X_j)' S^{-1} (X_i - X_j)} \tag{C1}$$

where X_i is a vector of observable characteristics for a parish belonging to a test district. In our case, these are average income; net wealth; employment shares in manufacturing and agriculture; population density; proportion of fertile married women; and a dummy variable for urban locations. S denotes the covariance matrix of the vector of observable characteristics. Since the matching was done before the data collection took place, it does not take our outcome variables into account. This is a virtue insofar as our matching procedure is based on information similar to that available to the decision makers at the time of the intervention. The matching was done in random order and without replacement. Further information on the identification of the control group and the underlying matching procedure is given in Bhalotra et al. (2017).

Table C1 shows descriptive statistics and standardised differences in means for the main outcome variables in our treatment and control regions in the pre-intervention period. None of the outcome variables appears to be unbalanced according to the standardised difference. The small differences that are present should also not be problematic since we control for parish specific differences and parish specific time trends.

Since we rely on difference-in-differences techniques to estimate the effect of the intervention it is desirable that a) it can be argued that the control group and the treatment group would

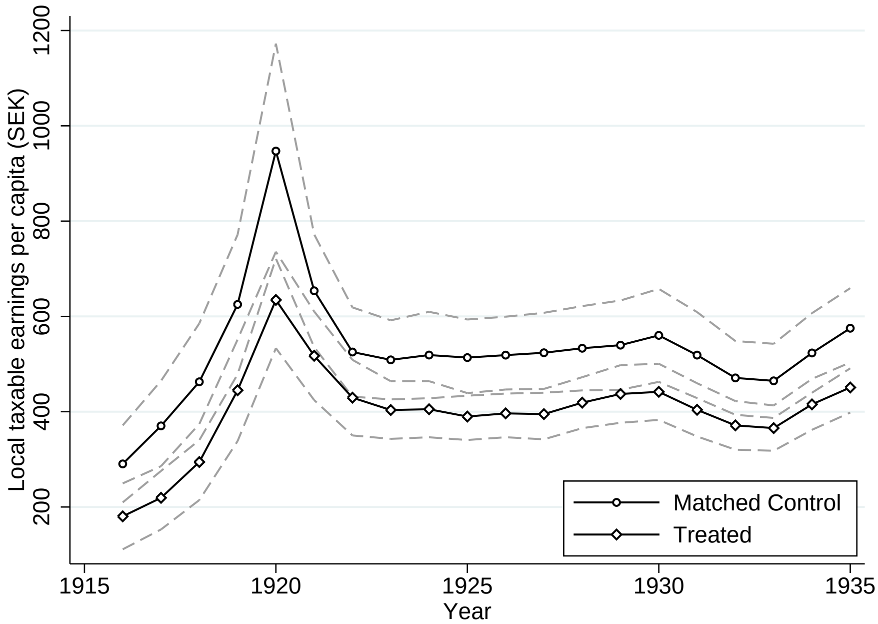
Table C1. Pre-Intervention Balance of Outcomes

	Control					Treated					
	Count	Mean	SD	Min	Max	Count	Mean	SD	Min	Max	Std Diff.
School Data											
Share Sickn. Abs.	1,923	0.043	0.062	0	1	2,290	0.051	0.073	0	1	-0.12389
Top GPA	1,935	0.160	0.366	0	1	2,353	0.210	0.407	0	1	-0.12867
GPA	1,935	3.510	0.702	1	6	2,353	3.519	0.674	1	6	-0.01297
Math	1,923	3.460	0.815	1	6	2,344	3.516	0.825	1	6	-0.06779
Reading/Speaking	1,931	3.567	0.768	1	7	2,349	3.556	0.711	1	6	0.01555
Writing	1,541	3.460	0.817	1	7	2,005	3.419	0.771	1	6	0.05180
Census 1970											
Secondary Schooling	1,468	0.138	0.345	0	1	1,919	0.178	0.382	0	1	-0.11016
Working Fulltime	1,456	0.650	0.477	0	1	1,904	0.641	0.480	0	1	0.02018
Working Parttime	1,456	0.130	0.337	0	1	1,904	0.137	0.344	0	1	-0.01934
Top Income	1,470	0.202	0.402	0	1	1,920	0.215	0.411	0	1	-0.03215
log Income	1,470	9.554	1.113	0	12	1,920	9.565	1.134	0	12	-0.00923
Municipal	1,456	0.167	0.373	0	1	1,904	0.148	0.355	0	1	0.05157
Governmental	1,456	0.079	0.270	0	1	1,904	0.085	0.279	0	1	-0.02223

have followed a common trend in the absence of the intervention; b) there were no contemporary and relevant changes that asymmetrically affected treated areas, and c) the intervention had no indirect impact on the local labour market. In Figure C1 we plot the average taxable earnings in treated and control areas. Even though they have been matched based on 1930 earnings according to the 1930 census, treated municipalities appear to be slightly poorer on average in terms of *taxable* earnings (although the confidence intervals overlap in all years). However, treated and control areas exhibit parallel trends throughout the 1917–35 period and there is no indication that the intervention had a contemporary effect on local earnings.

D Appendix: Utilisation

Table D1 exploits detailed utilisation data measured at the individual level to explore whether the gender driven effects could also be due to the uptake of utilisation for female children. The data stems from nurse and physician records archived for four of the seven health districts and covers a representative sample for about half of the eligible children (Bhalotra et al., 2017). We regress uptake of utilisation on duration of eligibility in years and interact this with a female dummy. Column 1 reports results for a linear model taking into account the number of visits, column 2 estimates a linear probability model with enrolment as a binary indicator and column 3 estimates utilisation conditional on enrolment and thus the intensive margin. As can be seen from the table, eligibility in years is a good predictor for utilisation but there is no higher



Note: Observations from control group were weighted based on their population size relative to the population size of treated locations they were matched to. Peak in 1920 due to inflation.

Figure C1. Local Taxable Earnings Per Capita.

uptake for female children. Thus, the gender specific effects are not due to gender differences in utilisation.

Table D1. Utilisation.

	OLS (1)	**LPM** (2)	**Cond. on Enrolment** (3)
Duration of Eligibility	3.0086***	0.5394***	2.6508***
	(0.823)	(0.043)	(0.894)
Female Child	-0.0008	-0.0045	0.4962
	(0.116)	(0.049)	(0.308)
Female×Duration of Eligibility	0.4126	0.0194	-0.0388
	(0.367)	(0.059)	(0.739)
In-Wedlock Birth	0.3880	-0.0042	0.8829
	(0.397)	(0.035)	(0.748)
Twin Birth	0.1363	0.0993	-0.5591
	(0.392)	(0.067)	(0.720)
Born to Younger Mother	-0.0427	0.0272	-0.2209
	(0.383)	(0.045)	(0.479)
Born to Older Mother	-0.0084	-0.0125	0.0732
	(0.105)	(0.026)	(0.248)
High SES	0.4427*	0.0063	0.8628*
	(0.251)	(0.029)	(0.514)
Low SES	-0.2232	-0.0815	0.2107
	(0.374)	(0.068)	(0.538)
Constant	-0.3754	0.1340**	1.1748**
	(0.417)	(0.054)	(0.517)
N	2,577	2,577	1,214
r2	0.052	0.138	0.018

Note: *** p <0,01; ** p <0,05; * p <0,1, Standard errors are clustered at the parish level. Outcome variable is uptake of utilisation of infant care.

E Appendix: Heterogeneity

Since the programme was especially targeted at vulnerable groups like children and mothers with a relatively disadvantaged background, we also conduct heterogeneity analyses to explore whether the intervention was beneficial for children born out of wedlock and those born to families of low socio-economic status. Children born out of wedlock were of special concern since they had significantly worse health prospects than other children during that time period (Stenhoff, 1931). Table E1 shows heterogeneity results for standardised marks and standardised GPA in fourth grade. Children born to single mothers experienced larger improvements in fourth grade regarding their GPA, 'reading and speaking' and 'writing' marks. This effect is mainly driven by males born to single mothers (results not shown here). We do not find any significant heterogeneity in treatment effects for long-term outcomes or in the first grade. The improvement in marks is between 0.17 and 0.4 standard deviations if they were eligible to the intervention one more year. This effect is relatively large and three to four times the magnitude of what we find for the whole population. This is in line with the findings of Bhalotra et al. (2017) who also find significant improvements in the reduction of mortality for children born out of wedlock. Although not significant, effects for children born into a low SES environment still point to an improvement in academic performance.

Table E1. Heterogeneity in treatment effects academic performance grade 4.

	GPA		**Math**		**Reading**		**Writing**	
	Single (1)	**low SES** (2)	**Single** (3)	**low SES** (4)	**Single** (5)	**low SES** (6)	**Single** (7)	**low SES** (8)
DID×*Variable*	0.1679	-0.0500	0.1853	-0.0123	0.0544	-0.0285	0.2733	-0.1086
	(0.143)	(0.077)	(0.170)	(0.102)	(0.168)	(0.095)	(0.173)	(0.087)
DID	0.0760*	0.1069**	-0.0182	0.0017	0.1230***	0.1381**	0.1230*	0.1813***
	(0.039)	(0.041)	(0.048)	(0.067)	(0.046)	(0.056)	(0.065)	(0.065)
Treated×*Variable*	-0.0138	0.0380	-0.0048	0.0369	0.0636	0.0069	-0.1027	0.0703
	(0.085)	(0.059)	(0.112)	(0.067)	(0.093)	(0.078)	(0.104)	(0.054)
Parish FE	✓	✓	✓	✓	✓	✓	✓	✓
QOB×YOB Effects	✓	✓	✓	✓	✓	✓	✓	✓
School FE	✓	✓	✓	✓	✓	✓	✓	✓
SES Effects	✓	✓	✓	✓	✓	✓	✓	✓
Length of Schoolyear	✓	✓	✓	✓	✓	✓	✓	✓
Schoolform	✓	✓	✓	✓	✓	✓	✓	✓
Parish Specific Linear Trends								
N	13,268	13,268	13,242	13,242	13,223	13,223	13,228	13,228
Pre-mean	-0.047	-0.047	-0.027	-0.027	-0.056	-0.056	-0.057	-0.057

Note: *** p <0,01; ** p <0,05; * p <0,1, Standard errors are clustered at the parish level. Covariates which are included in all specifications are a dummy indicating twin births, a dummy for being female, dummies capturing old (>35 years) and young (<20) mothers, a dummy for married women, a dummy indicating a hospital birth and the treatment effect of the maternal intervention. *Pre-mean* refers to the mean value of the outcome variable. *QOB× YOB effects* include quarter-of-birth dummies for each of the 20 quarters. *Parish FE* are fixed effects for the parish the individual lived in at the time of the birth. *SES effects* are fixed effects for the professional group of the household head. *Length of schoolyear* are fixed effects controlling for the reforms concerning the length of the school year. *Schoolform* are fixed effects controlling for the school form as described in Section 2.2 and *Parish×Birth date effects* allows for parish specific time trends. *Variable* refers to single mother respectively low SES mother interaction.

F Appendix: Attribution of Treatment Effects

F.1 Definitions

We now derive the attribution of treatment effects that has been conducted in Section 5.1. Consider two binary outcome variables W and Z for which, without loss of generality, the outcome 1 represents the "better" outcome. For example, we may think of W as representing secondary schooling completion and Z representing employment. We assume that individuals are exposed to a binary treatment D with associated treatment effects $\tau_{Wi} \equiv W_i^1 - W_i^0$ and $\tau_{Zi} \equiv Z_i^1 - Z_i^0$, where W_i^j and Z_i^j represents the potential outcome associated with treatment assignment j. Clearly, $\tau_{Wi} \in \{-1, 0, 1\}$ and $\tau_{Zi} \in \{-1, 0, 1\}$.[47]

We define the average treatment effects on the treated as $\tau_W = \mathbb{E}\left[\tau_{Wi} \mid D = 1\right]$ and $\tau_Z = \mathbb{E}\left[\tau_{Zi} \mid D = 1\right]$. Clearly, we have $\tau_W \in [-1, 1]$ and $\tau_Z \in [-1, 1]$. We would like to find out the extent to which the treatment effects are correlated so that individuals who positively contribute to τ_W (i.e. the individuals who have $\tau_{Wi} = 1$) are the same as the individuals who contribute positively to τ_Z (i.e. the individuals who have $\tau_{Zi} = 1$). In order to do so, we partition the population into 16 distinct groups, depending on their values for τ_{Wi} and τ_{Zi}.

Definition 1 (Population Partition) *Denote by $p_{kl,mn}$ the proportion of the treated population characterised by $(W^0 = k, W^1 = l, Z^0 = m, Z^1 = n)$. Thus,*

$$p_{kl,mn} = \Pr\left(W^0 = k, W^1 = l, Z^0 = m, Z^1 = n \mid D = 1\right) \forall (k, l, m, n) \in \{0, 1\}^4 . \qquad \text{(F1)}$$

Also, denote by p_{kl}^W the proportion of the treated population characterised by $(W^0 = k, W^1 = l)$ and by p_{mn}^Z the proportion of the treated population characterised by $(Z^0 = m, Z^1 = n)$.

For example, the subpopulation $p_{01,01}$ consists of individuals who have $\tau_{Wi} = \tau_{Zi} = 1$ and thus experience an increase in both outcomes when treated. With the population partition, we can define explicitly the average treatment effects:

$$\tau_W = p_{01,00} + p_{01,01} + p_{01,10} + p_{01,11} - p_{10,00} - p_{10,01} - p_{10,10} - p_{10,11} \qquad \text{(F2)}$$

$$\tau_Z = p_{00,01} + p_{01,01} + p_{10,01} + p_{11,01} - p_{00,10} - p_{01,10} - p_{10,10} - p_{11,10} \qquad \text{(F3)}$$

[47] We have defined the treatment indicator as binary here, although a small part of our sample has treatment exposure between zero and one. It would be possible to generalise this exposition to allow for non-binary treatment by imposing the assumption that effects are linear in treatment exposure.

This composition of τ_W (and analogously for τ_Z) follows immediately from the definition $\tau_W = \mathbb{E}\left[W_i^1 - W_i^0 \mid D = 1\right]$: the subpopulations with $W^0 = 0, W^1 = 1$ have size $p_{01,00} + p_{01,01} + p_{01,10} + p_{01,11}$ and the subpopulations with $W^0 = 1, W^1 = 0$ have size $p_{01,11} + p_{10,00} + p_{10,01} + p_{10,10} + p_{10,11}$. The other subpopulations are either always- or nevertakers who do not contribute to the estimated effect.

We may also derive the covariance of the two treatment effects:

$$\operatorname{Cov}\left(\tau_{Wi}, \tau_{Zi} \mid D = 1\right) = \mathbb{E}\left[\tau_{Wi} \cdot \tau_{Zi} \mid D = 1\right] - \tau_W \cdot \tau_Z \tag{F4}$$

$$= p_{01,01} + p_{10,10} - p_{01,10} - p_{10,01} - \tau_W \cdot \tau_Z \tag{F5}$$

Next, we move on to define the benchmark value of correlation between τ_{Wi} and τ_{Zi} which would obtain if the two treatment effects are completely unrelated so that $p_{kl,mn} = p_{kl}^W p_{mn}^Z \forall k, l, m, n \in \{0,1\}^4$

Definition 2 (Uncorrelated Effects) *We refer to the effects τ_{Wi} and τ_{Zi} as* ***uncorrelated*** *whenever $p_{kl,mn} = p_{kl}^W p_{mn}^Z \forall k, l, m, n \in \{0,1\}^4$ so that* $\operatorname{Cov}\left(\tau_{Wi}, \tau_{Zi}\right) = 0$

F.2 Implementation

Our implementation is based on defining a third outcome by interacting the two main outcomes. We thus introduce the new outcome variable $Y = W \cdot Z$ with associated treatment effect τ_{Yi} and average treatment effect τ_Y. We will now show that the value of τ_Y is well-defined under the assumption of uncorrelated effects and that the actual estimate for τ_Y can be compared to this benchmark in order to gauge how strongly correlated the effects are. In addition, we derive how the degree of correlation between treatment effects can be assessed under different assumptions regarding the underlying population.

Analogously to the definitions for τ_W and τ_Z above, the average treatment effect on Y is defined as follows:

$$\tau_Y = p_{01,01} + p_{01,11} + p_{11,01} - p_{10,10} - p_{11,10} - p_{10,11} \tag{F6}$$

The first three terms represent individuals who increase their value of the interacted outcome Y due to the treatment: they do so either if they increase both outcomes, or because they are compliers in one outcome and always-takers in the other. We now evaluate the treatment effects under the assumption that the treatment effects are uncorrelated.

Uncorrelated effects: In this scenario, the treatment effects on the two outcomes are completely uncorrelated and the estimate for τ_Y simplifies as follows:

$$\begin{aligned}
\tau_Y^{uc} &= p_{01,01} + p_{01,11} + p_{11,01} - p_{10,10} - p_{11,10} - p_{10,11} \\
&= p_{01}^W p_{01}^Z + p_{01}^W p_{11}^Z + p_{11}^W p_{01}^Z - p_{10}^W p_{10}^Z - p_{11}^W p_{10}^Z - p_{10}^W p_{11}^Z \\
&= \tau_W \tau_Z + \tau_W \Pr\left(Z^0 = 1\right) + \tau_Z \Pr\left(W^0 = 1\right)
\end{aligned}$$

For moderately-sized effects, this combined effect will be lower than the individual effects τ_W and τ_Z.[48] Besides, it is estimable, since $\Pr\left(Z^0 = 1\right)$ and $\Pr\left(W^0 = 1\right)$ are the missing counterfactuals that we impute using the common time trend.

We have thus shown that the uncorrelated case defined above translates into a clear benchmark value for τ_Y, which we denote τ_Y^{uc}. A comparison of τ_Y to this benchmark reveals whether treatment effects are correlated or not. Next, we impose an additional assumption in order to derive direct estimates of the degree of correlation between effects.

Assumption 1 (No 'defiers') *For all subpopulations and outcomes, the treatment effect is non-negative:* $\tau_{Ki} \geq 0 \forall i, K \in \{W, Z\}$.

The implication of assumption 1 is that all subpopulations with 'defiers' have mass zero: $p_{10,10} = p_{10,11} = p_{10,01} = p_{10,00} = 0$ and $p_{11,10} = p_{01,10} = p_{00,10} = 0$. If assumption 1 is satisfied, the three average treatment effects simplify as follows:

$$\begin{aligned}
\tau_W &= p_{01,01} + p_{01,11} + p_{01,00} \\
\tau_Z &= p_{01,01} + p_{11,01} + p_{00,01} \\
\tau_Y &= p_{01,01} + p_{01,11} + p_{11,01} \qquad \text{(F7)}
\end{aligned}$$

Our interest is in the size of the subpopulation $p_{01,01}$, since according to equation (F5) the size of this subpopulation determines how strongly correlated the outcomes are. In order to solve the system (F7) above, we introduce the two parameters $a = \frac{p_{01,00}}{p_{01,11}}$ and $b = \frac{p_{00,01}}{p_{11,01}}$. These parameters thus determine the relative sizes of the relevant subpopulations – and they lead to the solution

$$p_{01,01} = \frac{\tau_Y - \frac{\tau_W}{1+a} - \frac{\tau_Z}{1+b}}{1 - \frac{1}{1+a} - \frac{1}{1+b}} \qquad \text{(F8)}$$

[48] It is not a lower bound, since the two treatment effects could be *negatively* correlated – in which case τ_Y would take on even lower values. In most applications this possibility can be ruled out, however.

Thus, equation (F8) provides an estimate for $p_{01,01}$ which is defined as long as $1 - ab \neq 0$. Our analysis in section 5.1 is based on finding reasonable values for a and b based on the baseline distribution of W and Z in the population: since we have $\Pr(W = 1) \approx 0.2$ for all the binary outcomes we consider, our central assumption is that $a = b = 4$. This corresponds to a situation where compliers are proportional to the relative sizes of their respective subgroups in the population: $a = \frac{p_{01,00}}{p_{01,11}} = \frac{\Pr(Z^0=0)}{\Pr(Z^0=1)}$. In order to assess the robustness of the findings, we also consider scenarios where the compliers who are nevertakers for the other outcome (i.e. $p_{01,00}$ and $p_{00,01}$) are strongly underrepresented ($a = b = 2$) or strongly overrepresented ($a = b = 8$).

F.3 Results

Results for females are presented in Table 11 in the main text. Here we present results for males, which are of less interest since no adult outcomes are improved by the intervention – thus making the issue of correlated effects redundant. Since the "no defier" assumption cannot be satisfied for this group (due to some estimated effects being negative) we do not report the correlation coefficients of the effects.

Table F1. Correlated Effects Males.

Outcome 1	Outcome 2	(1) τ_1	(2) τ_2	(3) τ_Y^{uc}	(4) τ_Y	(5)	(6) corr (τ_{1i}, τ_{2i})	(7)
Top GPA (PRIMARY)								
	Secondary	0.0258	-0.0189	0.0012	0.0111			
		(0.033)	(0.036)		(0.018)			
	High Occ	0.0243	-0.0569	-0.0026	0.0274			
		(0.033)	(0.036)		(0.024)			
	Top Income	0.0243	-0.0252	0.0011	0.0254*			
		(0.033)	(0.042)		(0.014)			
Secondary Schooling								
	High Occ	-0.0314	-0.0702**	-0.0181	-0.0307*			
		(0.028)	(0.032)		(0.018)			
	Top Income	-0.0314	-0.0418	-0.0126	-0.0184			
		(0.028)	(0.034)		(0.019)			
High Occupation								
	Top Income	-0.0700**	-0.0414	-0.0230	-0.0521*			
		(0.032)	(0.034)		(0.029)			

Note: τ_Y^{uc}: benchmark value, uncorrelated effects (see Appendix F for a derivation); τ_1: treatment effect outcome 1; τ_2: treatment effect outcome 2; τ_Y: joint treatment effect for interacted outcome 1×2; corr (τ_{1i}, τ_{2i}): Correlation coefficient between treatment effects (Bounds for alternative assumptions in square brackets; see Appendix F for a derivation).

G Appendix: Child Day Care Expansion

From the early 1940's there was political debate in Sweden over state provision of preschool child care. At the time there were a few day care institutions run by charity organisations or the community (SOU, 1972)) or informal care.[49] Since 1943 the state provided partial salary funding of educated personnel (Hammarlund, 1998; Westberg, 2015) in day care centres, but these grants covered only about 10 per cent of total costs. In March 1963, the national government considerably extended this, covering a larger share of total costs and funding the training of more child-carers. As a result, many municipalities built new day care centres and employed day carers in the mid 1960's. The new state subsidy also led to the municipality taking over governance of most of the day care centres previously run by charity organisations (Westberg, 2015). From 1963 to 1970 the number of day care places increased from 10,300 to 29,200. As there remained an excess demand for day care, from 1968, the government implemented an expansion of family day care (*familjedaghem*), involving a day carer look after smaller groups of children (1-4) in their home (Karlsson, 2002). From January 1, 1969, every municipality was eligible for funding as long as the regional Child Welfare Board approved each centre and the carers are employed by the municipality (SOU, 1972). From 1969 to 1972, the number of places in municipality governed family day care increased from 10,000 to 42,000 (SOU, 1972).

As described in Section 4 we gathered data on the number of childcare workers per woman of reproductive age in 1960 and 1970 by municipality to mirror the expansion of state-subsidised child care. Table G1 shows that we can reject that the childcare expansion was correlated with the infant care trial. Table G2 shows that growth in employment and earnings among women exposed to the infant intervention was not significantly different in municipalities where childcare expansion was above vs below the median.

[49] The first day care institutions were established in the mid-19th century as charity institutions, which provided supervision, pedagogical activities and health care, and primarily were available to children of working mothers (Ekstrand, 2000; Hatje, 1999; Broman, 1995).

Table G1. Placebo Estimates: Effect of Treatment on Childcare Expansion.

	Level CC Employment		**Δ Employment 1960–70**	
DID Estimate	-0.0455	-0.0760	-0.0464	-0.0724
	(0.035)	(0.047)	(0.035)	(0.049)
R^2	0.460	0.473	0.457	0.468
Parish FE	✓	✓	✓	✓
QOB×YOB FE	✓	✓	✓	✓
School FE	✓	✓	✓	✓
SES Effects	✓	✓	✓	✓
Parish Trends		✓		✓

Note: *** p <0,01; ** p <0,05; * p <0,1. The dependent variables are the proportion of females aged 30–40 working in public childcare in 1970 (columns 1–2) and the change in that proportion between 1960–70 (columns 3–4). Childcare employment is measured at the municipality level ($N = 1,037$) and each individual is assigned to 1970 municipalities based on their parish of birth.

Table G2. Effect Heterogeneity by Childcare Expansion (Females).

	Level CC Empl.		**Δ Empl. 1960–70**	
A. Top Income				
Main DID Estimate	0.0733***	0.0805***	0.0627*	0.0709**
	(0.027)	(0.027)	(0.032)	(0.034)
Fully Interacted	-0.0153	-0.0027	0.0010	0.0107
	(0.039)	(0.041)	(0.042)	(0.045)
B. Ln Income				
Main DID Estimate	0.1775*	0.2274**	0.2168**	0.2792***
	(0.094)	(0.102)	(0.100)	(0.103)
Fully Interacted	-0.1178	-0.0698	-0.1708	-0.1421
	(0.111)	(0.113)	(0.115)	(0.118)
C. Full-time work				
Main DID Estimate	0.0701**	0.0775**	0.0751*	0.0834**
	(0.032)	(0.034)	(0.039)	(0.042)
Fully Interacted	-0.0198	-0.0040	-0.0256	-0.0103
	(0.050)	(0.052)	(0.053)	(0.055)
Parish FE	✓	✓	✓	✓
QOB×YOB FE	✓	✓	✓	✓
School FE	✓	✓	✓	✓
SES Effects	✓	✓	✓	✓
Parish Trends		✓		✓

Note: *** p <0,01; ** p <0,05; * p <0,1. The DID variable (Treated parish × Duration of eligibility) and its components are interacted with the proportion of females aged 30–40 working in public childcare in 1970 (columns 1–2) and the change in that proportion between 1960–70 (columns 3–4). Childcare employment is measured at the municipality level ($N = 1,037$) and each individual is assigned to 1970 municipalities based on their parish of birth. The parameter denoted "Fully Interacted" gives estimates for the interacte variable *Treated parish* × *Duration of eligibility* × *Childcare expansion.*

H Appendix: Robustness Checks

H.1 Selective Survival

To investigate the role of survival selection, we estimate programme effects on 1970 income for subsamples of individuals surviving until age 40, 50, 60, 70 and 75 respectively (see Appendix Table H1).

Table H1. Assessing survival selection: log earnings 1970 by age at observation

	Mean	**All** (1)	**Age 40** (2)	**Age 50** (3)	**Age 60** (4)	**Age 70** (5)	**Age 75** (6)
WOMEN							
log Earnings	8.990	0.1947***	0.1942***	0.2184***	0.2080***	0.2062***	0.2592***
		(0.066)	(0.068)	(0.065)	(0.068)	(0.071)	(0.076)
N		10,301	10,275	10,085	9,657	8,820	8,119
MEN							
log Earnings	10.222	-0.0464	-0.0459	-0.0377	-0.0532*	-0.0750**	-0.0921***
		(0.036)	(0.036)	(0.033)	(0.031)	(0.032)	(0.033)
N		10,619	10,574	10,177	9,408	8,041	7,006
Parish FE		✓	✓	✓	✓	✓	✓
QOB×YOB FE		✓	✓	✓	✓	✓	✓
SES Effects		✓	✓	✓	✓	✓	✓
School Reforms		✓	✓	✓	✓	✓	✓
Parish Trends		✓	✓	✓	✓	✓	✓

Note: *** p <0,01; ** p <0,05; * p <0,1, Standard errors are clustered at the parish level. Covariates which are included in all specifications are a dummy indicating twin births, a dummy for being female, dummies capturing old (>35 years) and young (<20) mothers, a dummy for married women, a dummy indicating a hospital birth and the treatment effect of the maternal intervention. *Mean* refers to the mean value of the outcome variable before the intervention took place. *QOB× YOB effects* include quarter-of-birth dummies for each of the 20 quarters. *Parish FE* are fixed effects for the parish the individual lived in at the time of the birth. *SES effects* are fixed effects for the professional group of the parental household head. *School reforms* refers to the extension of compulsory schooling and length of school year reforms and *Parish trends* allows for parish specific time trends.

During the time period of interest Sweden provided a large widow pension. We therefore control for a widow(er) dummy in Table H2. We also interact the widow(er) dummy with the partner's observed income in 1970 (since the pension was a function of the partner's earnings). According to these estimates, widows increase their incomes by 20 per cent on average, and it is proportional to their partner's income, suggesting we proxy it quite well. Importantly, our results are not affected by these inclusions.

H.2 Alternative Treatment Indicators

- Binary any exposure: =1 if duration for infant eligibility$>$ 0 days; 0 otherwise
- Binary min 3 months: =1 if eligible to infant intervention $\geq$ any 3 months in total; 0

Table H2. DID estimates pension age 71.

	log Pension Age 71 (1)	(2)	+Widow(er) Dummy (3)	(4)	+Widow(er) x Income Partner (5)	(6)
Pension Age 71						
DIDI	-0.0035	0.0187	-0.0069	0.0135	-0.0026	0.0159
	(0.012)	(0.014)	(0.014)	(0.013)	(0.013)	(0.013)
Widow			-0.0122	-0.0124	0.0138	0.0142
			(0.024)	(0.026)	(0.017)	(0.018)
Widow x income partner					0.0004	0.0003
					(0.002)	(0.003)
N	15,964	15,964	15,854	15,854	15,854	15,854
Pre-mean	11.789	11.789	11.789	11.789	11.789	11.789
Pension Age 71 Females						
DIDI	0.0293	0.0711***	0.0271	0.0654***	0.0331*	0.0688***
	(0.019)	(0.015)	(0.021)	(0.015)	(0.019)	(0.015)
Widow			0.2091***	0.2059***	0.0114	0.0116
			(0.009)	(0.011)	(0.038)	(0.039)
Widow x income partner					0.0051***	0.0050***
					(0.001)	(0.001)
N	8,284	8,284	8,225	8,225	8,225	8,225
Pre-mean	11.609	11.609	11.609	11.609	11.609	11.609
Pension Age 71 Males						
DIDI	-0.0400**	-0.0400*	-0.0447***	-0.0445**	-0.0422**	-0.0434**
	(0.017)	(0.020)	(0.017)	(0.020)	(0.016)	(0.020)
Widower			-0.0124	-0.0124	0.0136	0.0142
			(0.025)	(0.026)	(0.017)	(0.018)
Widower x income partner					0.0004	0.0003
					(0.002)	(0.002)
N	7,680	7,680	7,629	7,629	7,629	7,629
Pre-mean	11.995	11.995	11.995	11.995	11.995	11.995
Parish FE	✓	✓	✓	✓	✓	✓
QOB×YOB Effects	✓	✓	✓	✓	✓	✓
SES Effects	✓	✓	✓	✓	✓	✓
School Reforms	✓	✓	✓	✓	✓	✓
Parish Specific Linear Trends		✓		✓		✓

Note: *** p <0,01; ** p <0,05; * p <0,1, Standard errors are clustered at the parish level. Covariates which are included in all specifications are a dummy indicating twin births, a dummy for being female, dummies capturing old (>35 years) and young (<20) mothers, a dummy for married women, a dummy indicating a hospital birth and the treatment effect of the maternal intervention. *Pre-mean* refers to the mean value of the outcome variable before the intervention took place. *QOB× YOB effects* include quarter-of-birth dummies for each of the 20 quarters. *Parish FE* are fixed effects for the parish the individual lived in at the time of the birth. *SES effects* are fixed effects for the professional group of the parental household head. *School reforms* refers to the extension of compulsory schooling and length of school year reforms and *Parish specific linear trends* allows for parish specific time trends. Widow dummy is based on comparison of 1970 marital status and wedlock status during pension age. Income of partner in 1,000 SEK in 1970.

otherwise

- Binary at least first 3 months: =1 if birth date $\geq$ 1 October 1931 and birth date $\leq$ 31 March 1933; =2 if otherwise treated; 0 control group (omitted category).

- Binary 12 months/full eligibility: =1 if duration for infant eligibility=12 months; i.e. those born between 1 October 1931 and 30 June 1932; =2 if otherwise treated; 0 control group (omitted category).

Table H3. Results for different treatment indicators on long-term outcomes - Males.

	log Income 1970 (1)	**Working Parttime** (2)	**Working Fulltime** (3)	**Secondary Schooling** (4)	**Municipal** (5)	**Governmental** (6)
Binary Any Exposure	-0.0131	0.0005	-0.0003	-0.0482	0.0026	0.0086
	(0.031)	(0.008)	(0.015)	(0.034)	(0.019)	(0.016)
Binary Min 3 Months	-0.0289	-0.0023	-0.0061	-0.0239	-0.0123	0.0073
	(0.028)	(0.007)	(0.012)	(0.023)	(0.011)	(0.015)
Binary at Least First 3 Months Complete	0.0302	-0.0105	0.0282	-0.0311*	0.0226	-0.0021
	(0.040)	(0.009)	(0.019)	(0.019)	(0.027)	(0.024)
Binary Other Treated	-0.0376	0.0066	-0.0159	-0.0561	-0.0107	0.0170
	(0.038)	(0.010)	(0.018)	(0.048)	(0.017)	(0.015)
Binary 12 Months/Full Eligibility	0.0041	-0.0076	0.0083	-0.0260	0.0250	-0.0279
	(0.049)	(0.010)	(0.022)	(0.019)	(0.033)	(0.019)
Binary Other Treated	-0.0200	0.0034	-0.0033	-0.0549	-0.0074	0.0239
	(0.034)	(0.010)	(0.015)	(0.043)	(0.016)	(0.017)
N	10,619	10,466	10,466	10,613	10,466	10,466
Pre-mean	10.222	0.019	0.925	0.172	0.092	0.111
Parish FE	✓	✓	✓	✓	✓	✓
QOB×YOB Effects	✓	✓	✓	✓	✓	✓
SES Effects	✓	✓	✓	✓	✓	✓
School Reforms	✓	✓	✓	✓	✓	✓
Parish Specific Linear Trends	✓	✓	✓	✓	✓	✓

Note: *** p <0,01; ** p <0,05; * p <0,1, Standard errors are clustered at the parish level. Covariates which are included in all specifications are a dummy indicating twin births, a dummy for being female, dummies capturing old (>35 years) and young (<20) mothers, a dummy for married women, a dummy indicating a hospital birth and the treatment effect of the maternal intervention. *Pre-mean* refers to the mean value of the outcome variable and *mean income* lists average income in each sector by gender. *QOB× YOB effects* include quarter-of-birth dummies for each of the 20 quarters. *Parish FE* are fixed effects for the parish the individual lived in at the time of the birth. *SES effects* are fixed effects for the professional group of the parental household head. *School reforms* refers to the extension of compulsory schooling and length of school year reforms and *Parish specific linear trends* allows for parish specific time trends.

Table H4. Results for different treatment indicators on long-term outcomes - Females.

	log Income 1970	**Working Parttime**	**Working Fulltime**	**Secondary Schooling**	**Municipal**	**Governmental**
	(1)	(2)	(3)	(4)	(5)	(6)
Binary Any Exposure	0.0808	-0.0264	0.0397*	0.0192	0.0183	0.0156
	(0.075)	(0.029)	(0.023)	(0.015)	(0.017)	(0.010)
Binary Min 3 Months	0.1752***	-0.0394	0.0650*	0.0340*	0.0332**	0.0166*
	(0.053)	(0.037)	(0.033)	(0.019)	(0.016)	(0.009)
Binary at Least First 3 Months Complete	0.1408*	-0.0702	0.0966**	0.0341	0.0374	0.0250*
	(0.078)	(0.052)	(0.040)	(0.021)	(0.025)	(0.015)
Binary Other Treated	0.0481	-0.0032	0.0090	0.0117	0.0078	0.0105
	(0.088)	(0.022)	(0.022)	(0.019)	(0.020)	(0.010)
Binary 12 Months/Full Eligibility	0.0972	-0.0506	0.0874**	0.0364*	0.0371	0.0335**
	(0.090)	(0.036)	(0.037)	(0.020)	(0.025)	(0.015)
Binary Other Treated	0.0741	-0.0177	0.0220	0.0133	0.0111	0.0090
	(0.076)	(0.028)	(0.021)	(0.018)	(0.019)	(0.010)
N	10,301	10,256	10,256	10,297	10,256	10,256
Pre-mean	8.990	0.265	0.370	0.198	0.238	0.051
Parish FE	✓	✓	✓	✓	✓	✓
QOB×YOB Effects	✓	✓	✓	✓	✓	✓
SES Effects	✓	✓	✓	✓	✓	✓
School Reforms	✓	✓	✓	✓	✓	✓
Parish Specific Linear Trends	✓	✓	✓	✓	✓	✓

Note: *** p <0,01; ** p <0,05; * p <0,1, Standard errors are clustered at the parish level. Covariates which are included in all specifications are a dummy indicating twin births, a dummy for being female, dummies capturing old (>35 years) and young (<20) mothers, a dummy for married women, a dummy indicating a hospital birth and the treatment effect of the maternal intervention. *Pre-mean* refers to the mean value of the outcome variable and *mean income* lists average income in each sector by gender. *QOB× YOB effects* include quarter-of-birth dummies for each of the 20 quarters. *Parish FE* are fixed effects for the parish the individual lived in at the time of the birth. *SES effects* are fixed effects for the professional group of the parental household head. *School reforms* refers to the extension of compulsory schooling and length of school year reforms and *Parish specific linear trends* allows for parish specific time trends.

Table H5. Anchoring.

	Grade 1			Grade 4		
	Math (1)	**Reading** (2)	**Writing** (3)	**Math** (4)	**Reading** (5)	**Writing** (6)
1 Point	-1.196***	-0.696***	-1.107***	-0.319	0.483	-0.681***
	(0.263)	(0.262)	(0.325)	(0.252)	(0.761)	(0.254)
2 Points	-0.705***	-0.620***	-0.645***	-0.245***	-0.231	-0.331***
	(0.099)	(0.105)	(0.112)	(0.087)	(0.164)	(0.087)
4 Points	0.296***	0.288***	0.275***	0.282***	0.228***	0.185***
	(0.050)	(0.048)	(0.070)	(0.049)	(0.049)	(0.048)
5 Points	0.306*	0.518***	0.552**	0.522***	0.493***	0.415***
	(0.170)	(0.164)	(0.254)	(0.059)	(0.060)	(0.061)
6/7 Points	-0.837	0.363	1.051	0.638***	0.783***	0.533***
	(0.676)	(0.500)	(1.065)	(0.129)	(0.182)	(0.173)
Constant	9.865***	9.840***	9.870***	9.729***	9.722***	9.819***
	(0.025)	(0.026)	(0.030)	(0.037)	(0.040)	(0.035)
R^2	0.010	0.008	0.008	0.011	0.007	0.008
N	12,343	12,356	8,430	12,479	12,458	12,463

Note: SE in parenthesis, *** p <0,01; ** p <0,05; * p <0,1. Outcome variable is log income. The point values refer to the 7-point grading scale defined in Section 2.2 and 3. Reference group category 3. Marks 6 and 7 pooled due to few observations.

Table H6. DID estimates for subjects.

		Anchored Grading Scale			
		Math (1)	**Reading** (2)	**Writing** (3)	**Religion** (4)
Males and Females	**Grade 1 and 4**	-0.0004	0.0175**	0.0306***	0.0028
	SE	(0.010)	(0.008)	(0.012)	(0.009)
	N	26,403	26,400	22,234	26,298
	Pre-mean	9.907	9.913	9.906	9.909
Males and Females	**Grade 1**	-0.0038	0.0103	0.0255	-0.0001
	SE	(0.016)	(0.014)	(0.022)	(0.014)
	N	13,161	13,177	9,007	13,060
	Pre-mean	9.878	9.894	9.867	9.888
	Grade 4	-0.0012	0.0264**	0.0286**	0.0015
	SE	(0.013)	(0.010)	(0.012)	(0.013)
	N	13,242	13,223	13,227	13,238
	Pre-mean	9.934	9.932	9.933	9.928
Males	**Grade 1**	-0.0183	0.0201	0.0147	0.0062
	SE	(0.015)	(0.019)	(0.019)	(0.014)
	N	6,779	6,794	4,608	6,723
	Pre-mean	9.871	9.871	9.833	9.880
	Grade 4	0.0176	0.0365**	0.0379**	-0.0011
	SE	(0.021)	(0.015)	(0.015)	(0.027)
	N	6,688	6,687	6,692	6,689
	Pre-mean	9.919	9.893	9.884	9.891
Females	**Grade 1**	0.0105	0.0070	0.0398	-0.0039
	SE	(0.028)	(0.015)	(0.030)	(0.020)
	N	6,382	6,383	4,399	6,337
	Pre-mean	9.886	9.916	9.898	9.896
	Grade 4	-0.0171	0.0225*	0.0178	0.0088
	SE	(0.013)	(0.013)	(0.016)	(0.014)
	N	6,554	6,536	6,535	6,549
	Pre-mean	9.949	9.969	9.979	9.964
	Parish FE	✓	✓	✓	✓
	QOB×YOB Effects	✓	✓	✓	✓
	School FE	✓	✓	✓	✓
	SES Effects	✓	✓	✓	✓
	Length of Schoolyear	✓	✓	✓	✓
	Schoolform	✓	✓	✓	✓
	Parish Specific Linear Trends				

Note: *** p <0,01; ** p <0,05; * p <0,1, Standard errors are clustered at the parish-grade level. Covariates which are included in all specifications are a dummy indicating twin births, a dummy for being female, dummies capturing old (>35 years) and young (<20) mothers, a dummy for married women, a dummy indicating a hospital birth and the treatment effect of the maternal intervention. *Pre-mean* refers to the mean value of the outcome variable before the intervention took place. *QOB× YOB effects* include quarter-of-birth dummies for each of the 20 quarters. *Parish FE* are fixed effects for the parish the individual lived in at the time of the birth. *SES effects* are fixed effects for the professional group of the household head. *Length of schoolyear* are fixed effects controlling for the reforms concerning the length of the school year. *Schoolform* are fixed effects controlling for the school form as described in Section 2.2 and *Parish specific linear trends* allows for parish specific time trends.

Table H7. DID pre-trend test primary school.

	GPA_cog Stand. Grade 4	Math Stand. Grade 4	Writing Stand. Grade 4	Reading Stand. Grade 4	Share Sickn. Abs. Grade 4
	(1)	(2)	(3)	(4)	(5)
Females					
Interaction	0.0003	0.0009	0.0004	-0.0002	0.0000
	0.001	0.001	0.001	0.001	0.000
Trend	-0.0007*	-0.0011**	-0.0010**	0.0001	0.0000
	0.000	0.000	0.000	0.000	0.000
Treated	-0.1292	-0.1222	-0.1629	-0.1069	0.0040
	0.087	0.108	0.103	0.097	0.009
Constant	0.2558***	0.2069***	0.3591***	0.2033***	0.0440***
	0.063	0.078	0.074	0.070	0.006
N	1,050	1,048	1,048	1,049	1,048
Males					
Interaction	-0.0002	-0.0001	-0.0007	-0.0000	-0.0000
	0.001	0.001	0.001	0.001	0.000
Trend	-0.0007*	-0.0011**	-0.0003	-0.0007	0.0000
	0.000	0.000	0.000	0.000	0.000
Treated	0.0065	0.0100	0.0822	-0.0682	0.0131**
	0.087	0.107	0.104	0.100	0.006
Constant	-0.0792	0.0972	-0.2206***	-0.1191	0.0310***
	0.063	0.077	0.075	0.073	0.004
N	1,116	1,110	1,116	1,115	1,105

Note: *** p <0,01; ** p <0,05; * p <0,1, *Trend* variable is based on *month* × *year* observations; *Treated* refers to a dummy indicating treated parishes; *Interaction* is the interaction of the variables trend and treated.

Table H8. DID pre-trend test long-term outcomes.

	Secondary Schooling	Working Fulltime	Working Parttime	log Income 1970	Municipal	Governmental
	(1)	(2)	(3)	(4)	(5)	(6)
Females						
Interaction	-0.0000	0.0001	0.0000	-0.0000	-0.0000	0.0000
	0.000	0.000	0.000	0.001	0.000	0.000
Trend	0.0001	-0.0005**	-0.0000	-0.0010*	-0.0003	-0.0000
	0.000	0.000	0.000	0.001	0.000	0.000
Treated	0.0384	-0.0172	-0.0012	0.0154	-0.0345	-0.0022
	0.038	0.048	0.044	0.116	0.042	0.023
Constant	0.1307***	0.4277***	0.2526***	9.0252***	0.2894***	0.0555***
	0.028	0.036	0.033	0.086	0.031	0.017
N	1,665	1,656	1,656	1,666	1,656	1,656
Males						
Interaction	0.0002	-0.0000	0.0001	0.0000	-0.0001	0.0001
	0.000	0.000	0.000	0.000	0.000	0.000
Trend	-0.0002	0.0001	-0.0000	0.0001	0.0001	-0.0000
	0.000	0.000	0.000	0.000	0.000	0.000
Treated	0.0091	-0.0043	-0.0082	0.0408	0.0030	0.0009
	0.035	0.026	0.013	0.059	0.028	0.031
Constant	0.1589***	0.9198***	0.0184*	10.1674***	0.0851***	0.1079***
	0.027	0.020	0.010	0.045	0.021	0.024
N	1,722	1,704	1,704	1,724	1,704	1,704

Note: *** p <0,01; ** p <0,05; * p <0,1, *Trend* variable is based on *month* × *year* observations; *Treated* refers to a dummy indicating treated parishes; *Interaction* is the interaction of the variables trend and treated.

Table H9. Attrition in Sample and Subgroups

Sample	N	School Sample			Census 70 Sample			Pension Sample		
		Total	Death	Other	Total	Death	Other	Total	Death	Other
All	24,390	0.339	0.077	0.261	0.140	0.104	0.036	0.336	0.272	0.064
Treated Districts	13,802	0.358	0.081	0.277	0.145	0.108	0.037	0.333	0.269	0.064
Control Districts	10,588	0.318	0.073	0.245	0.135	0.099	0.036	0.339	0.275	0.063
Eligible Birthdates	13,490	0.339	0.078	0.260	0.143	0.107	0.036	0.340	0.281	0.059
Ineligible Birthdates	10,900	0.338	0.076	0.263	0.135	0.099	0.036	0.331	0.261	0.070
Treated and Eligible BirthDate	7,631	0.355	0.082	0.273	0.151	0.111	0.040	0.339	0.277	0.062
Control or Ineligible	16,759	0.332	0.075	0.257	0.136	0.101	0.035	0.335	0.270	0.064

Table H10. Adult Outcomes, School Sample.

	Males & Females				Females				Males			
	N	Mean	(1)	(2)	N	Mean	(3)	(4)	N	Mean	(5)	(6)
Top Income 1970	15,106	0.228	-0.0012	0.0193	7,468	0.245	0.0460	0.0714*	7,638	0.210	-0.0474	-0.0324
			(0.017)	(0.018)			(0.031)	(0.040)			(0.036)	(0.032)
log Income 1970	15,106	9.594	-0.0023	0.0595	7,468	8.990	0.0554	0.1621*	7,638	10.222	-0.0588	-0.0426
			(0.041)	(0.039)			(0.083)	(0.087)			(0.036)	(0.035)
Secondary	15,097	0.185	0.0004	0.0130	7,465	0.198	0.0471*	0.0594**	7,632	0.172	-0.0453	-0.0329
			(0.019)	(0.017)			(0.024)	(0.028)			(0.033)	(0.031)
log Pension Age 71	11,629	11.789	-0.0156	0.0109	6,061	11.609	0.0170	0.0712***	5,568	11.995	-0.0523**	-0.0578*
			(0.018)	(0.018)			(0.020)	(0.022)			(0.023)	(0.030)
Parish FE			✓	✓			✓	✓			✓	✓
QOB×YOB FE			✓	✓			✓	✓			✓	✓
SES Effects			✓	✓			✓	✓			✓	✓
School Reforms			✓	✓			✓	✓			✓	✓
Parish Trends				✓				✓				✓

Note: *** p <0,01; ** p <0,05; * p <0,1. The specifications used in this table correspond to those used in Tables 4 and 6 in Section 4; however, we drop from the analysis sample individuals who are not present in the school sample.

Table H11. Lee Bounds for Census 1970 & Pension Outcomes.

	Males & Females				Females				Males			
	N	Mean	(1)	(2)	N	Mean	(3)	(4)	N	Mean	(5)	(6)
SECONDARY SCHOOLING												
Upper Bound	20,806	0.178	-0.0020	0.0047	10,244	0.193	0.0513***	0.0508***	10,562	0.161	-0.0542*	-0.0409
			(0.019)	(0.016)			(0.016)	(0.015)			(0.032)	(0.028)
Estimate	20,908	0.185	-0.0063	0.0026	10,295	0.198	0.0350**	0.0347**	10,613	0.172	-0.0468	-0.0289
			(0.017)	(0.013)			(0.016)	(0.014)			(0.029)	(0.021)
Lower Bound	20,822	0.186	-0.0076	0.0012	10,255	0.200	0.0338**	0.0335**	10,567	0.172	-0.0483	-0.0306
			(0.018)	(0.013)			(0.016)	(0.014)			(0.029)	(0.021)
TOP INCOME 1970												
Upper Bound	20,827	0.224	0.0132	0.0231*	10,261	0.241	0.0740***	0.0860***	10,566	0.207	-0.0467	-0.0389
			(0.015)	(0.013)			(0.022)	(0.028)			(0.032)	(0.027)
Estimate	20,918	0.228	0.0098	0.0208	10,299	0.245	0.0653***	0.0787***	10,619	0.210	-0.0445	-0.0361
			(0.016)	(0.013)			(0.022)	(0.028)			(0.034)	(0.028)
Lower Bound	20,819	0.229	0.0083	0.0195	10,256	0.246	0.0625***	0.0752***	10,563	0.211	-0.0450	-0.0355
			(0.016)	(0.013)			(0.022)	(0.028)			(0.033)	(0.028)
LOG INCOME 1970												
Upper Bound	20,808	9.584	0.0477	0.0907***	10,245	8.976	0.1547**	0.2285***	10,558	10.216	-0.0571	-0.0453
			(0.031)	(0.028)			(0.060)	(0.065)			(0.036)	(0.036)
Estimate	20,918	9.594	0.0292	0.0730**	10,299	8.990	0.1199*	0.1943***	10,619	10.222	-0.0596	-0.0464
			(0.032)	(0.028)			(0.063)	(0.066)			(0.037)	(0.036)
Lower Bound	20,825	9.605	0.0060	0.0500*	10,260	9.005	0.0853	0.1635**	10,565	10.230	-0.0721**	-0.0624*
			(0.032)	(0.029)			(0.062)	(0.067)			(0.035)	(0.034)
PENSIONS AGE 71												
Upper Bound	15,787	11.776	0.0140	0.0289**	8,210	11.599	0.0456**	0.0802***	7,577	11.979	-0.0212	-0.0290*
			(0.013)	(0.012)			(0.018)	(0.015)			(0.016)	(0.017)
Estimate	15,963	11.789	-0.0035	0.0187	8,283	11.609	0.0293	0.0712***	7,680	11.995	-0.0400**	-0.0400*
			(0.012)	(0.014)			(0.019)	(0.015)			(0.017)	(0.020)
Lower Bound	15,789	11.808	-0.0170	0.0061	8,214	11.627	0.0244	0.0667***	7,575	12.014	-0.0630***	-0.0618***
			(0.012)	(0.015)			(0.017)	(0.017)			(0.016)	(0.019)
Parish FE			✓	✓			✓	✓			✓	✓
QOB×YOB FE			✓	✓			✓	✓			✓	✓
SES Effects			✓	✓			✓	✓			✓	✓
School Reforms			✓	✓			✓	✓			✓	✓
Parish Trends				✓				✓				✓

Note: *** p <0,01; ** p <0,05; * p <0,1. The specifications used in this table correspond to those used in Tables 4 and 6 in Section 4. Upper and lower bounds are estimated using Lee's method (Lee, 2009). The trimming was carried out conditioning on child gender, and SES and marital status of the household head.

Table H12. School Outcomes Grade 4

	Males & Females				Females				Males			
	N	Mean	(1)	(2)	N	Mean	(3)	(4)	N	Mean	(5)	(6)
TOP GPA												
Upper Bound	12,915	0.176	0.0766*	0.0861**	6,378	0.229	0.1029*	0.1255*	6,444	0.119	0.0514	0.0374
			(0.039)	(0.040)			(0.056)	(0.068)			(0.035)	(0.029)
Estimate	13,167	0.173	0.0693*	0.0801**	6,465	0.227	0.1000*	0.1243*	6,607	0.116	0.0400	0.0275
			(0.038)	(0.039)			(0.058)	(0.070)			(0.033)	(0.028)
Lower Bound	12,919	0.162	0.0496	0.0640*	6,363	0.214	0.0824	0.1093	6,462	0.108	0.0211	0.0132
			(0.033)	(0.036)			(0.054)	(0.068)			(0.027)	(0.025)
GPA												
Upper Bound	12,956	-0.027	0.1323***	0.1369***	6,397	0.118	0.1315**	0.1535**	6,464	-0.179	0.1434**	0.1232*
			(0.033)	(0.031)			(0.065)	(0.071)			(0.056)	(0.070)
Estimate	13,167	-0.047	0.0787**	0.0851**	6,465	0.098	0.0410	0.0617	6,607	-0.200	0.1213**	0.1084
			(0.033)	(0.036)			(0.048)	(0.053)			(0.056)	(0.070)
Lower Bound	12,819	-0.079	0.0333	0.0399	6,342	0.073	-0.0022	0.0217	6,381	-0.240	0.0643	0.0485
			(0.029)	(0.037)			(0.036)	(0.042)			(0.052)	(0.072)
Parish FE			✓	✓			✓	✓			✓	✓
QOB×YOB FE			✓	✓			✓	✓			✓	✓
School FE			✓	✓			✓	✓			✓	✓
SES Effects			✓	✓			✓	✓			✓	✓
Length of Schoolyear			✓	✓			✓	✓			✓	✓
Schoolform			✓	✓			✓	✓			✓	✓
Parish Trends				✓				✓				✓

Note: *** p <0,01; ** p <0,05; * p <0,1. The specifications used in this table correspond to those used in Table 3 in Section 4. Upper and lower bounds are estimated using Lee's method (Lee, 2009). The trimming was carried out conditioning on child gender, and SES and marital status of the household head.

I Appendix: Cost-Benefit Calculations

The trial was fully funded by the central Government and the cost at the time of implementation was \$139,380 (SEK 41,400). Our estimates show that the intervention led to an approximately 19 per cent increase in earnings for women, with no corresponding gains for men. We now calculate the net present value of these earnings gains, assuming they were constant in proportional terms over the life cycle. We conduct the calculation based on two different sources: first, using the baseline 1970 earnings of untreated women, assuming that earnings were flat over the life cycle. Second, we use earnings by age, sex and year reported in Johansson et al. (2006), based on administrative records. This alternative measure provides a more faithful representation of the age profile of earnings of these cohorts. Since no gains in earnings were estimated for males, we exclude them from this analysis, and measure female benefits in proportion to total costs.[50]

To calculate the discounted benefits according to the first source (i.e. our analysis sample), we use the DID counterfactual for earnings of treated women in the absence of treatment: $y_{11}^0 = y_{10} + [y_{10} - y_{00}] = 8.86 + [8.85 - 8.82]$. Thus, the estimate for earnings in the absence of the intervention would be SEK $7,259$ which corresponds to $\$6,309$ in 2018 USD. We assume these earnings apply at all ages from 21–64 and calculate the NPV as:

$$NPV(Y) = Y_i \cdot \sum_{a=21}^{64} \frac{1}{(1+r)^a} \tag{I1}$$

where $Y_i = 6,309$, and r is the discount rate. Since 3,180 females were eligible for treatment, and since the average treatment intensity was 0.64 in that group, we need to multiply these NPVs with $0.19 \times 3,180 \times 0.64 = 387$ to get the total benefit in the population. Table I13 presents the results. The second column shows the NPV of lifetime earnings of an average eligible woman.

[50] According to historical sources, the bulk of costs accruing were fixed and common for the two arms of the intervention (prenatal and infant care). Therefore we consider the total costs of the intervention as a whole (SOU, 1935).

Table I13. NPV of Benefits (2018 USD)

Disc. Rate	NPV	Benefit
0.206	360	139,380
0.050	38,289	14,818,000
0.025	98,353	38,063,000

Notes: Earnings were converted into 2018 SEK using (Edvinsson, 2016) and then converted to 2018 USD using an exchange rate of 9.1. *Source:* Census 1970. Own calculations.

The internal rate of return according to this calculation is 0.206. As an alternative, we use actual earnings of women by age and year from Johansson et al. (2006). This publication reports total earnings (including capital earnings). In the year 1970 total earnings are reported to be $10,714$ in the 35–39 age group. Thus we adjust the series by factor $0.68 = \frac{7{,}259}{10{,}714}$ in order to make it compatible with the earnings of untreated women in our sample.[51] Our calculation of NPV or earnings now becomes

$$NPV(Y) = \sum_{a=18}^{64} \frac{Y_{ai}}{(1+r)^a} \tag{I2}$$

where Y_{ai} is the earnings reported for women of age a in the year they reach that age. The corresponding NPVs are given in Table I14.

Table I14. NPV of Benefits (2018 USD).

Disc. Rate	NPV	Benefit
0.187	360	139,380
0.050	38,894	15,052,000
0.025	117,468	45,460,000

Notes: Earnings were converted into 2018 SEK using (Edvinsson, 2016) and then converted to 2018 USD using an exchange rate of 9.1. *Source:* Johansson et al. (2006). Own calculations.

[51] The reason this adjustment is necessary is that the definition of income in Johansson et al. (2006) is broader and also includes e.g. capital gains.

The numbers derived using the two different approaches are largely consistent with each other: in both cases, we find that the intervention generated net benefits an order of magnitude larger than the costs. The internal rate of return equals 0.188 according to this alternative source.

J Appendix: Figures, Tables and Graphs

Table J1. School form

	Form	Sample	1940/1941
Full Time Attendance	A	37.42%	44.9%
	B1	33.81%	26.4%
	B2	18.25%	19.2%
	B3	2.53%	3.3%
	D1	7.16%	2.5%
	aid-class	-	1.4%
Half Time Attendance	C	0.53%	2.1%
	D2	0.31%	0.2%
	D3	-	0.0%

Note: Folkskolan was divided into Main forms (A, B1, B2, B3, and aid class) and Exception forms (C, D2 and D3). The main forms required full time reading and a teacher with an appropriate teacher degree, while the exception form included either half-time reading schools or schools with a teacher that did not have an appropriate teacher degree. The exception forms were only accepted if the local conditions allowed for no other forms. The table shows the proportion of school forms in our sample in comparison to official statistics for the school year 1940/41 (SOU, 1944).

Table J2. Descriptive statistics explanatory variables.

	All Live Births N=24,390		**School Data N=16,089**	**Census 1970 N=20,921**	**Pension Age71 N=16,194**
	Mean	SD	Mean	Mean	Mean
Female	0.485	0.500	0.493	0.492	0.524
Wedlock	0.895	0.307	0.921	0.902	0.907
Twin	0.026	0.160	0.023	0.023	0.023
Treated	0.566	0.496	0.551	0.565	0.569
Mother<20	0.052	0.222	0.044	0.050	0.049
Mother>35	0.226	0.418	0.238	0.223	0.222
hospital birth	0.295	0.456	0.253	0.298	0.302
SES Professional/Technical	0.029	0.168	0.023	0.029	0.029
SES administrative/Managerial	0.024	0.153	0.020	0.024	0.025
SES Clerical	0.015	0.121	0.013	0.015	0.015
SES Sales	0.026	0.158	0.023	0.025	0.026
SES Service	0.027	0.163	0.019	0.026	0.025
SES Agricultural	0.381	0.486	0.405	0.386	0.393
SES Production	0.397	0.489	0.410	0.399	0.394
SES Unknown	0.101	0.301	0.087	0.095	0.091
DurationI	0.353	0.402	0.354	0.351	0.351
DurationM	0.257	0.315	0.256	0.257	0.260
Duration	0.610	0.587	0.610	0.608	0.610

Note: Variable descriptions for the variables in this table are available in Appendix B.

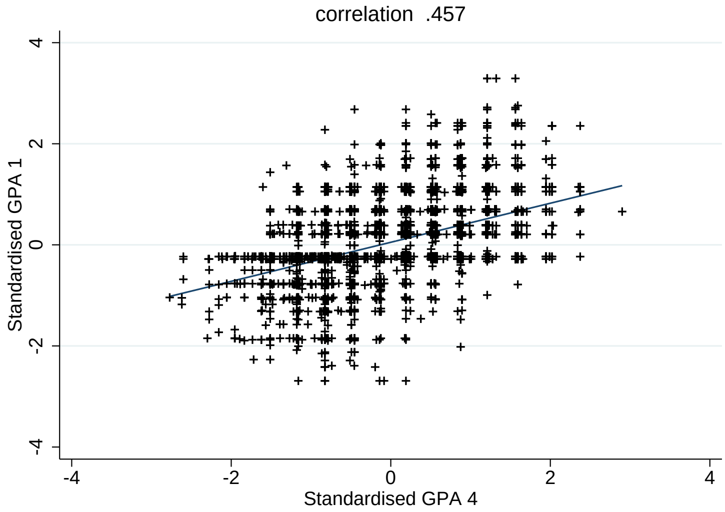


Figure J1. Correlation GPA in grade 1 and grade 4.

Table J3. Occupational Sorting: Disaggregated Results

	Men & Women (N=20,920)				Women (N=10,301)				Men (N=10,619)			
	Mean				Mean				Mean			
	Outc.	*Earn.*	(1)	(2)	*Outc.*	*Earn.*	(3)	(4)	*Outc.*	*Earn.*	(5)	(6)
A1. Technical	0.050	39,326	0.0154	0.0026	0.004	24,633	0.0022	-0.0009	0.097	39,906	0.0282	0.0060
			(0.014)	(0.008)			(0.003)	(0.003)			(0.028)	(0.016)
A2. Health care	0.047	19,403	0.0163**	0.0149**	0.084	18,650	0.0324**	0.0272**	0.008	28,613	0.0006	0.0029
			(0.008)	(0.006)			(0.015)	(0.012)			(0.006)	(0.006)
A3. Pedagogical	0.061	38,271	-0.0125	0.0010	0.069	30,822	-0.0031	0.0135	0.052	47,020	-0.0218	-0.0114
			(0.010)	(0.007)			(0.011)	(0.010)			(0.014)	(0.009)
B1. Bookkeeping	0.021	21,794	-0.0105*	-0.0038	0.031	18,565	-0.0096	-0.0005	0.011	33,581	-0.0113	-0.0072
			(0.006)	(0.006)			(0.009)	(0.010)			(0.009)	(0.006)
B2. Office work	0.059	22,612	0.0225**	0.0152*	0.093	18,917	0.0484*	0.0447*	0.025	32,841	-0.0029	-0.0139
			(0.011)	(0.009)			(0.026)	(0.023)			(0.009)	(0.013)
Parish FE			✓	✓			✓	✓			✓	✓
QOB×YOB FE			✓	✓			✓	✓			✓	✓
SES Effects			✓	✓			✓	✓			✓	✓
School Reforms			✓	✓			✓	✓			✓	✓
Parish Trends				✓				✓				✓

Note: *** p <0,01; ** p <0,05; * p <0,1, Standard errors are clustered at the parish level. Letters A and B refer to the main categories in Table 9. We provide means of the dependent variables as shares of men and women working in the occupational category at baseline (Outc.) and also mean earnings for each occupation (Earn.). Covariates which are included in all specifications are a dummy indicating twin births, a dummy for being female, dummies capturing old (>35 years) and young (<20) mothers, a dummy for married women, a dummy indicating a hospital birth and the treatment effect of the maternal intervention. *Mean* refers to the mean value of the outcome variable before the intervention took place. *QOB×YOB effects* include quarter-of-birth dummies for each of the 20 quarters. *Parish FE* are fixed effects for the parish the individual lived in at the time of the birth. *SES effects* are fixed effects for the professional group of the parental household head. *School reforms* refers to the extension of compulsory schooling and length of school year reforms and *Parish trends* allows for parish specific time trends.

Table J4. Gelbach Mediation – Females

	Secondary Schooling			**High Occupation**			**Earnings**		
Treatment Effect			0.0484*			0.0605			0.1861**
SE			(0.027)			(0.057)			(0.091)
N			6,105			6,105			6,105
Pre-mean			0.189			0.318			9.036
Unexplained = Treatment Effect - $\hat{\delta}$			0.0164			0.0378			0.0938
	$\hat{\Gamma}$	$\hat{\beta}$	$\hat{\delta} = \hat{\Gamma} \times \hat{\beta}$	$\hat{\Gamma}$	$\hat{\beta}$	$\hat{\delta} = \hat{\Gamma} \times \hat{\beta}$	$\hat{\Gamma}$	$\hat{\beta}$	$\hat{\delta} = \hat{\Gamma} \times \hat{\beta}$
Top GPA	0.1073*	0.2951***	0.0320*	0.1084*	0.0642***	0.0070***	0.1084*	0.0398	0.0043
	(0.063)	(0.016)	(0.018)	(0.063)	(0.020)	(0.003)	(0.063)	(0.033)	(0.005)
Secondary Schooling				0.0484*	0.3236***	0.0157*	0.0484*	0.2029***	0.0098
				(0.027)	(0.027)	(0.009)	(0.027)	(0.067)	(0.007)
High Occupation							0.0605	1.2922***	0.0782
							(0.057)	(0.046)	(0.070)

Note: $\hat{\beta}$ refers to estimates from full model of interest (dependent variable see columns); $\hat{\Gamma}$ refers to estimates from auxiliary models with each possible mediator acting as dependent variable; $\hat{\delta}$ is component of omitted variable bias estimated to be due to each variable (see Gelbach, 2016). Reference category gpa: 1st quintile.

Table J5. Gelbach Mediation Males.

	Secondary Schooling			High Occupation			Earnings		
Treatment Effect			-0.0243			-0.0636*			-0.0388
SE			(0.036)			(0.035)			(0.039)
N			6,121			6,121			6,121
Pre-mean			0.149			0.262			10.214
Unexplained = Treatment Effect - $\hat{\delta}$			-0.0243			-0.0618			-0.0108
	$\hat{\Gamma}$	$\hat{\beta}$	$\hat{\delta} = \hat{\Gamma} \times \hat{\beta}$	$\hat{\Gamma}$	$\hat{\beta}$	$\hat{\delta} = \hat{\Gamma} \times \hat{\beta}$	$\hat{\Gamma}$	$\hat{\beta}$	$\hat{\delta} = \hat{\Gamma} \times \hat{\beta}$
Top GPA	0.0213	0.3236***	0.0074	0.0229	0.0459	0.0010	0.0229	0.0844	0.0019
	(0.034)	(0.039)	(0.011)	(0.034)	(0.031)	(0.002)	(0.034)	(0.053)	(0.003)
Secondary Schooling				-0.0243	0.4941***	-0.0120	-0.0243	0.1174*	-0.0029
				(0.036)	(0.023)	(0.017)	(0.036)	(0.060)	(0.005)
Top GPA & No Secondary				0.0122	0.0862**	0.0010	0.0122	-0.0475	-0.0006
				(0.023)	(0.042)	(0.002)	(0.023)	(0.053)	(0.001)
High Occupation							-0.0636*	0.5107***	-0.0325**
							(0.035)	(0.060)	(0.016)
High Occ & No Secondary							-0.0326	-0.2043***	0.0067
							(0.031)	(0.051)	(0.006)

Note: $\hat{\beta}$ refers to estimates from full model of interest (dependent variable see columns); $\hat{\Gamma}$ refers to estimates from auxiliary models with each possible mediator acting as dependent variable; $\hat{\delta}$ is component of omitted variable bias estimated to be due to each variable (see Gelbach, 2016).